



Polarization-Based Haze Removal Using Self-Supervised Network

Yingjie Shi^{1,2}, Enlai Guo^{1,2*}, Lianfa Bai^{1,2} and Jing Han^{1,2*}

¹Jiangsu Key Laboratory of Spectral Imaging and Intelligent Sense, Nanjing University of Science and Technology, Nanjing, China,

²School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China

Atmospheric scattering caused by suspended particles in the air severely degrades the scene radiance. This paper proposes a method to remove haze by using a neural network that combines scene polarization information. The neural network is self-supervised and online globally optimization can be achieved by using the atmospheric transmission model and gradient descent. Therefore, the proposed method does not require any haze-free image as the constraint for neural network training. The proposed approach is far superior to supervised algorithms in the performance of dehazing and is highly robust to the scene. It is proved that this method can significantly improve the contrast of the original image, and the detailed information of the scene can be effectively enhanced.

OPEN ACCESS

Edited by:

Haofeng Hu,
Tianjin University, China

Reviewed by:

Jing Guan,
North University of China, China
Abhishek K. Singh,
The Ohio State University,
United States

*Correspondence:

Enlai Guo
njustgel@njust.edu.cn
Jing Han
eohj@njust.edu.cn

Specialty section:

This article was submitted to
Optics and Photonics,
a section of the journal
Frontiers in Physics

Received: 04 October 2021

Accepted: 22 November 2021

Published: 03 January 2022

Citation:

Shi Y, Guo E, Bai L and Han J (2022)
Polarization-Based Haze Removal
Using Self-Supervised Network.
Front. Phys. 9:789232.
doi: 10.3389/fphy.2021.789232

Keywords: polarization, dehazing, neural network, selfsupervised, haze remove method

1 INTRODUCTION

The existence of haze, due to the tiny water droplets or solid particles suspended in the air, brings many inconveniences to daily life. The air can no longer be regarded as an isotropic medium which leads to scattering of the transmitted light. The scene image received by the camera or human eyes has a severe degradation. As the distance from the target increases or the concentration of suspended particles increases, the scattering becomes more and more serious. Therefore, the details of the distant target are more severely lost, and the contrast of the captured image is also reduced more. Eliminating the influence of haze on the collected image is often required which can make it easier for the observer to identify the target.

The current methods for dehazing mainly include the data-driven method [1–3], the method based on prior knowledge [4–6], and the method based on physical models [7–10]. The first two types of methods hardly contain physical models, therefore, the problem these methods solved is essentially ill-posed. Data-driven methods often need to obtain a large number of hazy-clean pairs in advance for training and use deep learning or image feature extraction methods to achieve haze removal. The method based on prior knowledge mainly combines some statistical characteristics in the image contained haze. Appropriate parameters need to be selected and combined with the prior model to remove haze in the acquired image. Most of these methods can achieve dehazing through one image, but the limited information contained in the single image cannot provide the unique characters of the scene. Changes of the scene or objects with special colors in the scene may cause the failure of dehazing [4, 11, 12]. As another type of approach, methods based on physical models can solve the shortcomings of the above two types of methods to a certain extent. Physical-model-based methods often use the depth map or analyze the changes in the polarization state of the scene. These methods often need to take multiple images, through the depth map of the scene or the polarization intensity difference, to obtain the transmission map during the scattering process. Both of these methods can construct a unique model based on the characteristics of the scene itself, so the haze can

be removed more accurately. But sometimes methods based on physical models also require empirical knowledge to select appropriate filtering parameters [13–15].

Many data-driven or prior knowledge-based methods have emerged in the field of computer vision to achieve haze removal. Cai et al. realized dehazing through a single frame image by an end-to-end structure Convolutional Neural Network (CNN) [1]. A total of 100,000 sets of data are used for the model training during the experiment; such a huge amount of data consumes a lot of time in the collection and calculation process. Akshay et al. used the Generative Adversarial Networks (GAN) to achieve dehazing with a single frame image. The simulation data is used in the training process, and this strategy causes the trained model to not be well applied to actual scenarios [16]. He et al. analyzed the color distribution in the haze image and proposed a Dark Channel Prior (DCP) method for dehazing. But this method may be invalid when the target color in the scene is inherently similar or close to the background airtight (such as a white wall, snowy ground, etc.) [4].

The earliest dehazing process often uses polarization information to build a physical model. Schechner et al. used the polarization state difference of the scene due to scattering to achieve haze removal. However, it is necessary to manually select the window in the picture to determine the airlight intensity, which will introduce a lot of errors [7]. In recent years, polarization-based methods have been continuously developed. Shen et al. proposed a dehazing method by using the polarization state information to iteratively find the transmission map [17]. Liu et al. used polarization to separate the high-frequency and low-frequency information of the scene to achieve dehazing [13]. Shen et al. used the fusion of polarization intensity, hue, and saturation to achieve dehazing [18]. Because scene information such as depth can be extracted from the polarization difference of the two frames scene image, these methods can be used in most scenes without a priori. These methods may need to adjust the angle of the polarizer to obtain the two images with the largest polarization difference, so the data collection process is cumbersome.

This paper proposes a Polarization-based Self-supervised Dehazing Network named PSDNet that combines the difference of polarization information with deep learning to eliminate the influence of haze on the image. The feature map of the neural network is activated through the transmission map calculated by the scene polarization state. Then the transmission map with more accurate depth information is estimated and has richer detail. The transmission map, haze-free image, and airlight can be calculated by the network and a self-supervised closed loop is formed to optimize the network. Because the physical model is used as a constraint, huge amounts of data are no longer needed to optimize the weight of the network. PSDNet only needs two frames of orthogonal polarization state images of the scene as input to remove scene haze based on online training. The global optimization of the neural network also solves the problem of inaccurate selection of airlight and makes it dehaze more accurately. Compared with similar methods, the proposed method can more effectively improve the visibility of target details and is highly robust to the scene.

2 METHOD AND MODEL DESIGN

2.1 Basics of Polarization-Based Dehazing

When imaging through the atmosphere containing haze, the particles in the atmosphere will cause scattering of the scene radiance which leads to degrading the target image. As shown in **Figure 1**, the scattered scene radiance and the scattered light from the illumination are received by the camera. The intensity of airlight increases as the distance increases, which can be expressed as

$$A = A_{\infty}(1 - t(x)), \quad (1)$$

where A_{∞} is the intensity of atmospheric light at infinity, and $t(x)$ represents the rate of transmission at position x , describing the scattering and absorption of radiance in the atmosphere, $t(x)$ is given by

$$t(x) = \exp\left(-\int_0^x \beta(x') dx'\right), \quad (2)$$

where $\beta(x')$ is the extinction coefficient caused by scattering or absorption. When the extinction coefficient in the atmosphere does not change with distance, $\beta(x') = \beta$, **Eq. 2** can be written as

$$t(x) = \exp(-\beta x). \quad (3)$$

The process of removing the haze from the image is to restore the radiance intensity and color information of the original scene, which is usually modeled as

$$I(x) = L(x)t(x) + A_{\infty}(1 - t(x)), \quad (4)$$

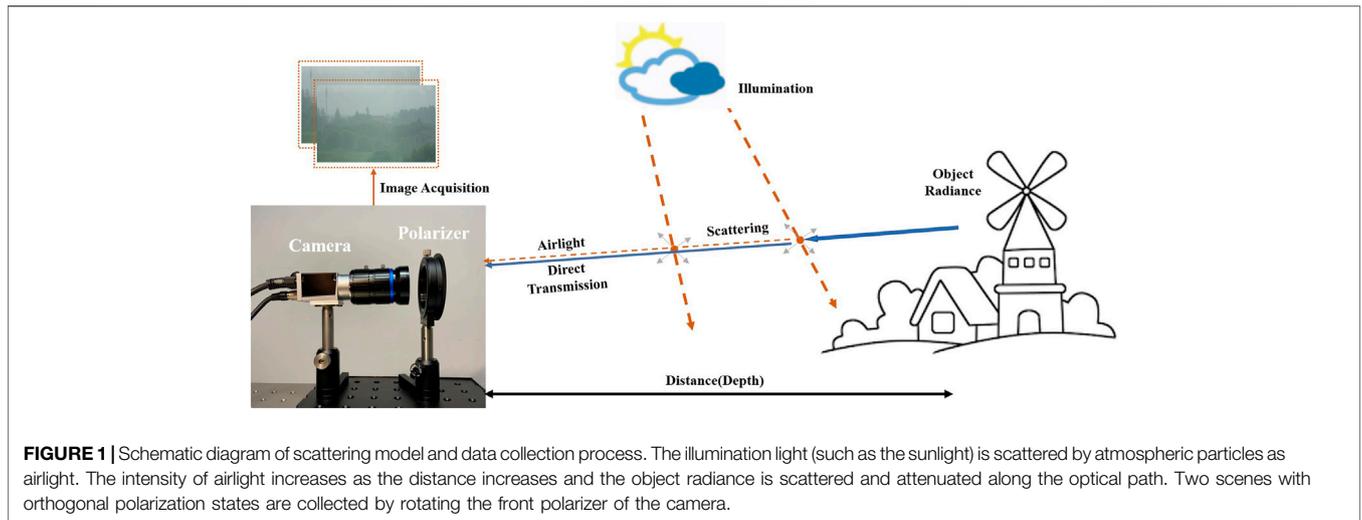
where $L(x)$ is the radiance of the scene at position x when there are no scattering particles in the atmosphere, and it is also the “clear image.” $L(x)$ can be expressed as

$$L(x) = \frac{I(x) - A_{\infty}(1 - t(x))}{t(x)}, \quad (5)$$

where $I(x)$ denotes the degraded version of $L(x)$ by atmospheric scattering. The effects of scattering on the polarization characteristics have been extensively studied. Generally, in the process of imaging through the atmosphere containing scattering particles, the degree of polarization of the original scene is almost negligible. The polarization is more related to the scattering process in the transmission of optical signals and is sensitive to the scattering distance [7]. Therefore, the transmission map can be calculated according to the difference in the polarization state in the captured image. A plane can be defined according to the light ray from the source to the scatterer and the line of sight from the camera. The airlight can be divided into two polarization components that are parallel and perpendicular to this plane, named A^{\parallel} and A^{\perp} respectively. The degree of polarization of airlight can be calculated by

$$p = \frac{A^{\perp} - A^{\parallel}}{A}, \quad (6)$$

where



$$A = A^\perp + A^\parallel, \tag{7}$$

is the total radiance due to airlight, and A also equal to $A_\infty(1 - t(x))$. The intensity of A^\parallel and A^\perp is related to the size of the scattering particles in the scene. In some published dehazing methods by using polarization, the parallel component is associated with the minimum measured radiance at a pixel and the perpendicular component is associated with the maximum radiance. This limitation requires rotating the polarizer during data collection to ensure that the two components have the largest difference, which increases the time for data collection. PSDNet only needs two images that have a polarization difference and has no limitation to the degree of polarization difference, so only two frames of orthogonal polarization scene images at any angle are needed. To avoid confusion in the calculation, stipulate that $A^\perp > A^\parallel$. The airlight at any point in the captured picture can be estimated by

$$A^{scene} = \frac{I^\perp - I^\parallel}{p} \tag{8}$$

where I^\perp and I^\parallel are the scene images taken when the polarization direction is the same as A^\perp and A^\parallel . The transmission map t is calculated by

$$t = 1 - \frac{A^{scene}}{A_\infty} = 1 - \frac{(I^\perp - I^\parallel)}{pA_\infty} = 1 - \frac{(I^\perp - I^\parallel)A}{(A^\perp - A^\parallel)A_\infty} \tag{9}$$

therefore, the airlight intensity at infinity A_∞ only needs to be estimated to recover the radiance of the scene without haze. The brightest point in the image is often considered as A_∞ . Although those strategies have good performance in most scenes, the brightest light intensity cannot accurately express A_∞ when white objects appear in the scene. The accuracy of manually selecting the A_∞ will also affect the final dehazing result. In addition, the reliability of the transmission map also determines the quality of the haze removed image, and the accuracy of the atmospheric degree of polarization also affects the accuracy of the transmission map. Airlight is generally

considered as partially linearly polarized light. With the rotation of the polarization axis of the polarizer, the rise and fall of the light intensity can be observed. The maximum and minimum light intensity are needed in the degree of polarization calculation, and if the polarization axis orientation of the polarizer cannot correspond to the direction of the airlight polarization, the degree of polarization is calculated inaccurately, which will cause the calculation error of the transmission map. Given the limitations of these methods, PSDNet is designed in which all calculations are in the same optimization process, so the transmission map and the airlight can be estimated simultaneously and accurately.

2.2 Model Design

To remove haze and get clear images, it is essential to obtain the transmission map and airlight, so PSDNet consists of three subnetworks, as shown in **Figure 7**. PSDNet-L, PSDNet-T, and PSDNet-A are used to calculate the target radiation L^{object} , transmission map, and the scene airlight A^{scene} respectively. Both PSDNet-L and PSDNet-T consist of convolution layers and pooling layers, and the structure of the network does not have a downsampling process which can reduce the loss of more detail. The last layer of all sub-networks uses the sigmoid function to normalize the output. Since the attributes of airlight are not related to the original scene distribution, PSDNet-A is composed of an encoder and a decoder, which are down-sampled and up-sampled respectively to extract global features and estimate the airlight [19].

The subnetwork PSDNet-T consists of two segments: PSDNet-T¹ and PSDNet-T². PSDNet-T¹ can extract the features of the scene from the haze image, then the transmission map estimated by the network is obtained. Meanwhile, the transmission map by using the conventional approach also can be calculated, and this calculation process uses the airlight estimated by PSDNet-A and the original image. When the scene polarization difference is minor as shown in **Figure 2B**, the calculated transmission map can respond to any

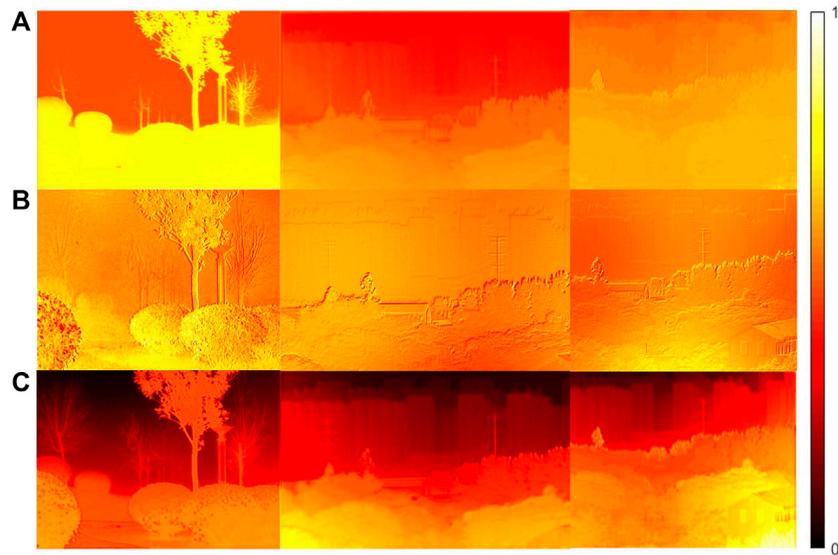


FIGURE 2 | Comparison of the transmission map. **(A)** Transmission map estimated by PSDNet- T^1 . **(B)** Transmission map calculated by scene polarization state. **(C)** Transmission map for the final use dehazing after fusion by PSDNet- T^2 .

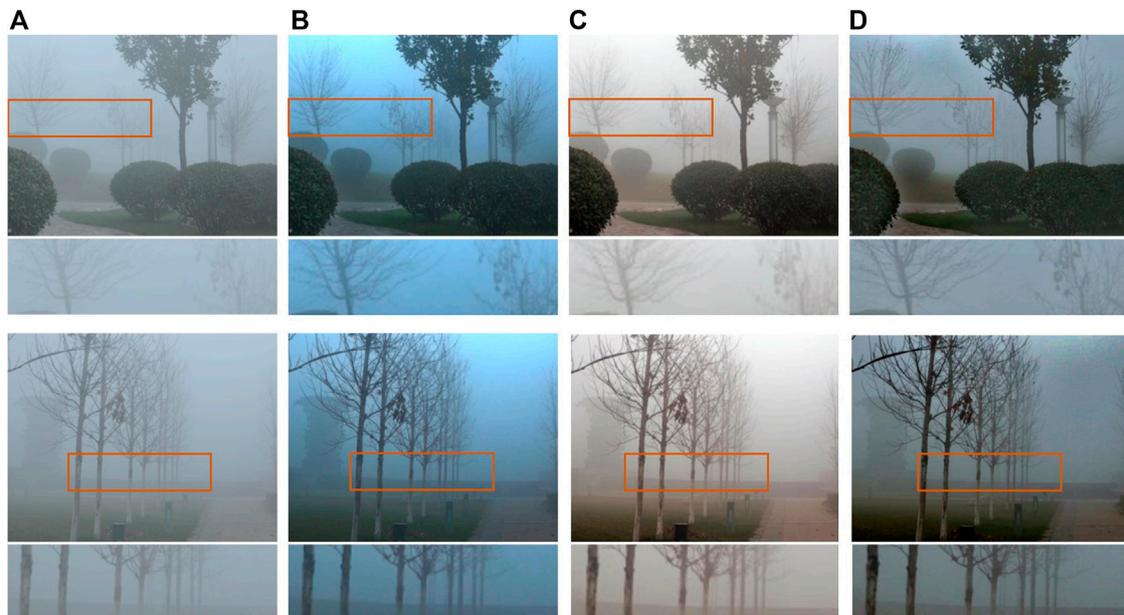


FIGURE 3 | The performance of different methods in the open-source dataset. **(A)** Original haze image. **(B–D)** Result of the DCP method, IIDWP method, and the proposed method respectively.

part of the scene but is often discontinuous. To be able to carry out effective dehazing, those transmission maps need to be properly filtered. The neural network can extract continuous feature maps as shown in **Figure 2A**, but details may be lost due to the lack of label constraints. Therefore, the PSDNet- T^2 is designed to fuse the feature maps and the transmission map calculated by polarization, and the transmission map involved in dehazing is obtained finally. In the fusion results as shown in

Figure 2C, the inaccurate rate of transmission is corrected, and the transmission map has a higher contrast which will enhance the final dehazing effect.

Finally, the clean image, transmission image, and atmospheric light estimated by the neural network are synthesized according to **Eq. 4**. The Mean Square Error (MSE) is used as a loss function to calculate the difference between the synthesized haze image and the real image. The MSE is formulated as:

$$MSE = \frac{1}{H*W} \sum_{i=1}^H \sum_{j=1}^W [I'(i, j) - I(i, j)]^2, \quad (10)$$

where I' is the synthesized haze image and the I is the image collected in the real scene. H and W are the height and width of those images, respectively. Different from the supervised algorithm, the self-supervised constraint strategy makes PSDNet not need a lot of haze-free images as the Ground Truth (GT) to constrain the optimization of the neural network. The results of dehazing depend on the quality of the transmission map and airlight. The effective use of polarization information makes it easy to estimate the transmission map more correctly, and the structure of the network combined with the physical model allows airlight to be estimated more accurately, then the original irradiance of the scene can be restored more effectively.

3 EXPERIMENT

3.1 Experimental Comparison on the Open-Source Dataset

An iterative image dehazing method with polarization (IIDWP) is proposed by Linghao Shen et al. [17]. Both the IIDWP method and the method proposed in this paper use the iterative optimization approach and scene polarization for dehazing. However, the IIDWP method only performs the iterative operation in the transmission map calculations process, and the final haze-free image quality may still be affected by airlight estimation or parameter selection. The method proposed in this paper is based on global learning optimization. And there is no need to set algorithm parameters; the airlight estimation and transmission map calculation are in the same iterative process, which makes it easier to optimize to the globally optimal result. Shen et al. provide an open-source dataset that contains haze images with orthogonal polarization states [17]. And this open-source dataset is utilized to compare dehazing performance among different methods firstly. A classic method using a single-frame for dehazing, the method based on the Dark Channel Prior (DCP) [4], is also selected as a comparison. The provided original image resolution is 942*609 pixels, and all images are resized as 960*576 pixels to facilitate convolution calculation in neural networks. In the comparison experiment, the dehaze results exposed by the author who provided the original data are used directly.

Two scenes with severe pattern degradation are selected for comparison, as shown in **Figure 3**. In these two selected scenes, detailed information such as the ends of branches is severely lost due to the high density of haze. In terms of increasing image contrast, all three dehazing methods work admirably; however, the results by using methods based on DCP or IIDWP have substantial color aberrations in the sky. Thanks to the global optimization strategy of the PSDNet, the optimal airlight and the corresponding transmission map can be estimated more accurately. Therefore, the proposed method not only can better enhance the scene details but also preserve the color information of the original image.

3.2 Comparison With Similar Approach on the Captured Dataset by Ourselves

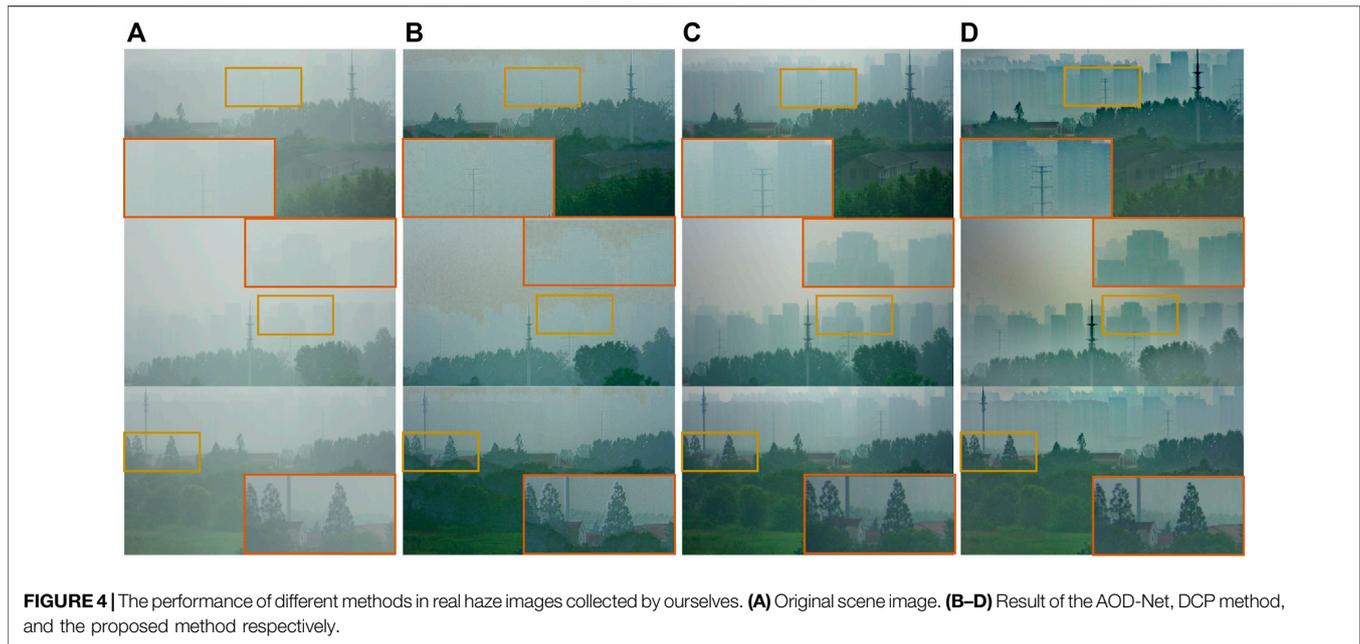
Although the proposed method is learning-based, it can perform self-supervision based on the polarization prior and physical model. So the PSDNet does not require GT as a constraint of the neural network compared with supervised networks. It is worth noting that Li et al. designed an end-to-end neural network named AOD-Net, which also incorporates the atmospheric transmission model [12]. As a representative of learning-based supervised algorithms for dehazing, the performance of AODNet is used as a comparison. In addition, the result of a method based on DCP is also used to compare the performance of the different approaches.

The nature of the supervised algorithm determines that AODNet requires a lot of data to build the association between haze images and haze-free images. It is difficult to collect massive hazy-clean pairs in the real scene, but the depth information of the picture is easier to obtain, so the dataset required for network training can be generated based on **Eq. 4, 1**. Haze image provided by NYU-Depth V2 [20] is simulated based on the depth images, which is the public dataset of New York University. Both the simulated haze dataset and the real outdoor dataset RESIDE-beta collected in Beijing [21] are used as the training dataset. A total of 50,000 hazy-clean pairs are used to train AOD-Net, and the other 10,000 pairs are used to verify the effectiveness of the trained model.

The data used for comparison was taken on a hazy morning, and the scenes are filmed from a distance of between 1 and 4 km. The system for pictures collection consists of a rotatable polarizer ($\phi = 50.8$ mm, extinction ratio = 1,000:1) and a color industrial camera (Basler, acA1920-40gc), and a telephoto industrial camera lens ($f = 100$ mm, 8 megapixels) is mounted on the camera. All original images have a raw resolution of 1920*1,200 pixels, and the center area with a size of 1920*1,156 pixels is cropped and rescaled to 960*576 pixels.

The final saved model is used to compare the dehazing effect of the method. The reference training epoch of AODNet is 40. To further improve the accuracy of the trained model, the final training epoch is increased to 50 and more than 36 h are used for training. In contrast, the proposed approach does not need to be trained in advance with the data mentioned above, only the haze image is needed as input and perform online learning. Therefore, the online training epoch of PSDNet is 800 but the consumed time is less than 5 min. And all the training environment is PyTorch 1.2.0 with RTX TITAN with I7-9700 CPU under ubuntu 16.04.

The scene image used to compare the effectiveness of different methods is captured in severe haze weather. The original scene image is shown in **Figure 4A**, and buildings in the distance need to be carefully discerned to see the outline, and the details are almost indistinguishable. **Figure 4B** shows the result of dehazing by the supervised algorithm AOD-Net. Although this method can effectively remove haze in the close-up of the scene, specifics about the distant scene are almost no enhancement. The main reason is that although huge amounts of data are used to train the network, these data cannot contain all scenarios in practice.



Ultimately, the trained models cannot be well applied to the widely varying real scenarios. **Figure 4C** shows the result of using the DCP to recover the original scene. The contour information in the long-distance can be distinguished after multiple parameter selection and tuning, but some details still cannot be recovered effectively. **Figure 4D** shows the dehazing result by using the proposed method, and haze removal can be more successful whether the scene is a close-distance or a long-distance. In the first scene, the windows on the buildings can be distinguished after dehazing by the proposed method, but it is completely sightless in the original image. And the tower crane in the zoomed-in area achieves visibility in the second scene, which is sightless too in the original.

The above comparison is almost intuitive; in terms of objective criteria, the result of image edge extraction can reflect the contrast level of the image. The edge extraction results of a high contrast image are more complete, and the target in the image is easier to distinguish. In the event of an image with low contrast, the opposite outcome is produced. Therefore, the dehazing results of different methods are subjected to edge detection to compare the image cleanliness from a more objective point of view. The Prewitt operator is a discrete differential operator which is often used in edge detection algorithms. At each point in the image, the result by using the Prewitt operator is either the corresponding gradient vector or the norm of the vector. The Prewitt operator is used to extract the edges of the dehazing results. Because the gradient approximation has a certain smoothing effect on the noise, edges cannot be extracted in low-contrast images, which is more conducive to contrast. The haze concentration in the far-field is much greater, and it is less visible in the original image without dehazing, so the detail and completeness of the edges extracted can reflect the quality of the dehazing result.

As shown in **Figure 5**, the original haze image and the dehazing result by AOD-Net can barely extract the edge contours of the distant buildings. Although the dehazing result by the DCP method can be detected to a certain extent, some distant building outlines are incomplete. The most complete edge of the distant contours can be extracted in the PSDNet dehazing results. The superior dehazing ability of PSDNet compared with other methods is shown, and the results of the comparison are also consistent with the visualization effect.

Since the proposed method does not require clean images as GT constrains neural networks, there are no haze-clean pairs used to assess haze removal quality. Therefore, the image quality assessment method that requires reference data cannot be used. But in order to analyze the ability of different methods to remove haze more objectively, contrast, saturation, and ENIQA [22] are selected as evaluation indexes to analyze the results of haze removal corresponding to different methods. Haze significantly reduces the contrast and saturation of the captured image, so for dehazing results, the higher the two indicators, the better the resolution of the target details. ENIQA is a high-performance general-purpose no-reference (NR) image quality assessment (IQA) method based on image entropy. The image features are extracted from two domains. In the spatial domain, the mutual information between the color channels and the two-dimensional entropy is calculated. In the frequency domain, the two-dimensional entropy and the mutual information of the filtered sub-band images are computed as the feature set of the input color image. Then the support vector machine is used to classify and give the indicator, and the final output score is between 0 and 1; the lower the score, the higher the image quality. In addition, different methods have great different dehazing abilities of different distance scene images, so the image is divided into two parts that is distant scene and the nearby scene in the objective indicators calculation process.

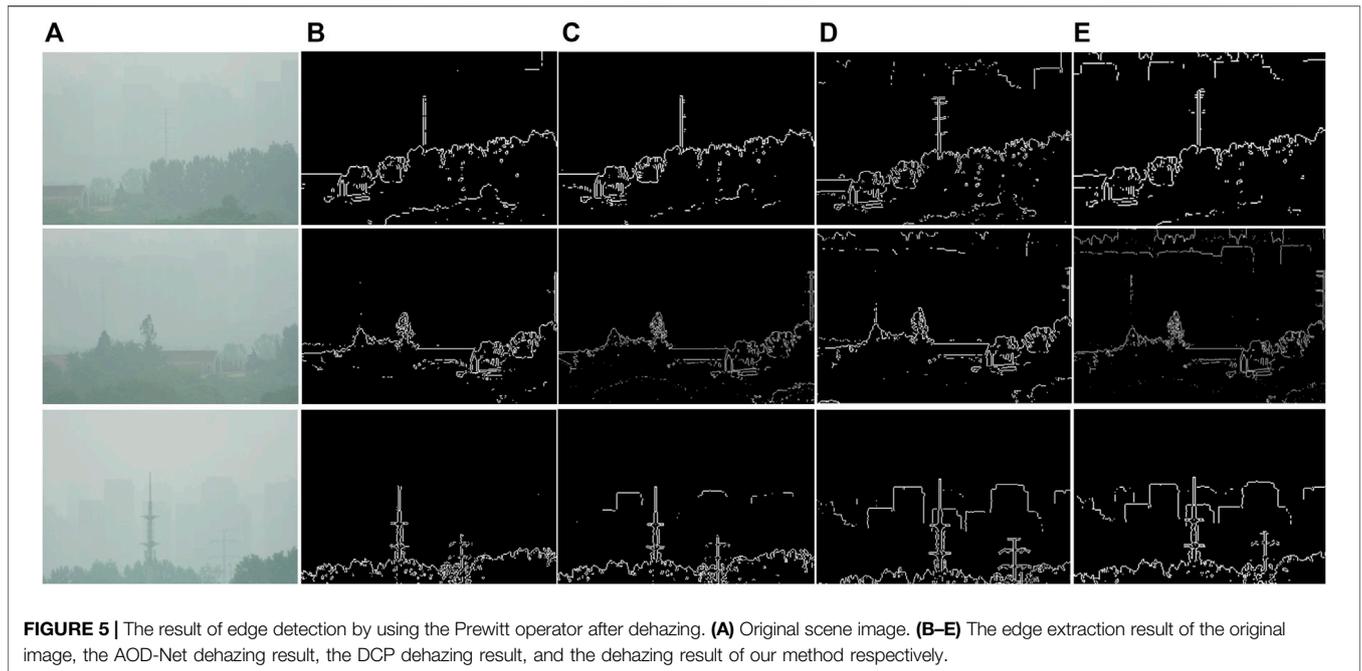


FIGURE 5 | The result of edge detection by using the Prewitt operator after dehazing. **(A)** Original scene image. **(B–E)** The edge extraction result of the original image, the AOD-Net dehazing result, the DCP dehazing result, and the dehazing result of our method respectively.

TABLE 1 | Indicator for the part of the image that contains a distant scene.

	Original	AODNet	DCP	Ours	IIDWP
Contrast	0.047	0.070	0.133	0.158	0.135
Saturation	0.074	0.083	0.093	0.106	0.089
ENIQA	0.392	0.381	0.391	0.274	0.280

Bold values indicate the best values.

TABLE 2 | Indicator for the part of the image that contains a nearby scene.

	Original	AODNet	DCP	Ours	IIDWP
Contrast	0.116	0.278	0.235	0.128	0.240
Saturation	0.076	0.224	0.179	0.138	0.163
ENIQA	0.392	0.190	0.285	0.137	0.239

Bold values indicate the best values.

The average haze removal indicator for the part of the picture that contains a distant scene is shown in **Table 1**, and the indicator for the part that contains a nearby scene is shown in **Table 2**. From the point of view of picture contrast and saturation index, the proposed method can be more effectively dehazing in the distant scene, and AODNet can be more effectively dehazing in the nearby scene; the conclusion is also consistent with subjective evaluation. This is mainly because AODNet conducts point-to-point optimization through haze-clean pairs in the training process, and the dehazing ability is limited to the scenes in the training set. And this approach cannot adapt to images with large differences in haze distribution compares to the training dataset. The proposed method utilizes the property of polarization changing of light during transmission in an atmosphere containing haze; therefore, a distant scene where

the light travels further can be used to estimate the transmission map more accurately, and the clearer details can be recovered. Besides, as the distance increases the effect of haze on image quality becomes more severe, the enhancement of detail in the image containing distant scenes is more useful. In addition, when ENIQA is used as evaluate indicator, the proposed method can improve the image quality in both scenes.

3.3 Robust and Efficiency of PSDNet

To demonstrate the robust of PSDNet, experiments, in which trained models of different scenes are used to remove the haze on one against another, are designed. Two types of scenes are selected in the training process, as shown in **Figure 6**; *scene 1* has similar distribution with *scene 3*, and *scene 2* has major differences with *scene 3*. Compared with *scene 3*, *scene 1* is collected on the same day and has the same haze distribution. In both scenes with trees in the near and buildings in the far distance, it should be noted that they are collected at different angles and the target distribution is not the same. *Scene 2* is the hazy polarization data disclosed by Shen et al. [17]. The two scenes of weather, target distribution, and illumination are different, moreover, *scene 2* is composed of plants and without buildings at a distance.

As shown in **Figure 6E**, when the model trained with *scene 1* is used as the pre-loaded model, the dehazing result of *scene 3* by direct inference in the detail improvement surpasses the result by using the method based on DCP. During this dehazing process, only the pre-loaded model is used, and PSDNet without any online training. Performance in the final comparison also reflects that the PSDNet combined with the physical model has good robustness to the different scenes. When the model trained with *scene 2* is used to remove the haze of *scene 2*, the result is shown in **Figure 6F**; only 25 iterations of online learning are required to get

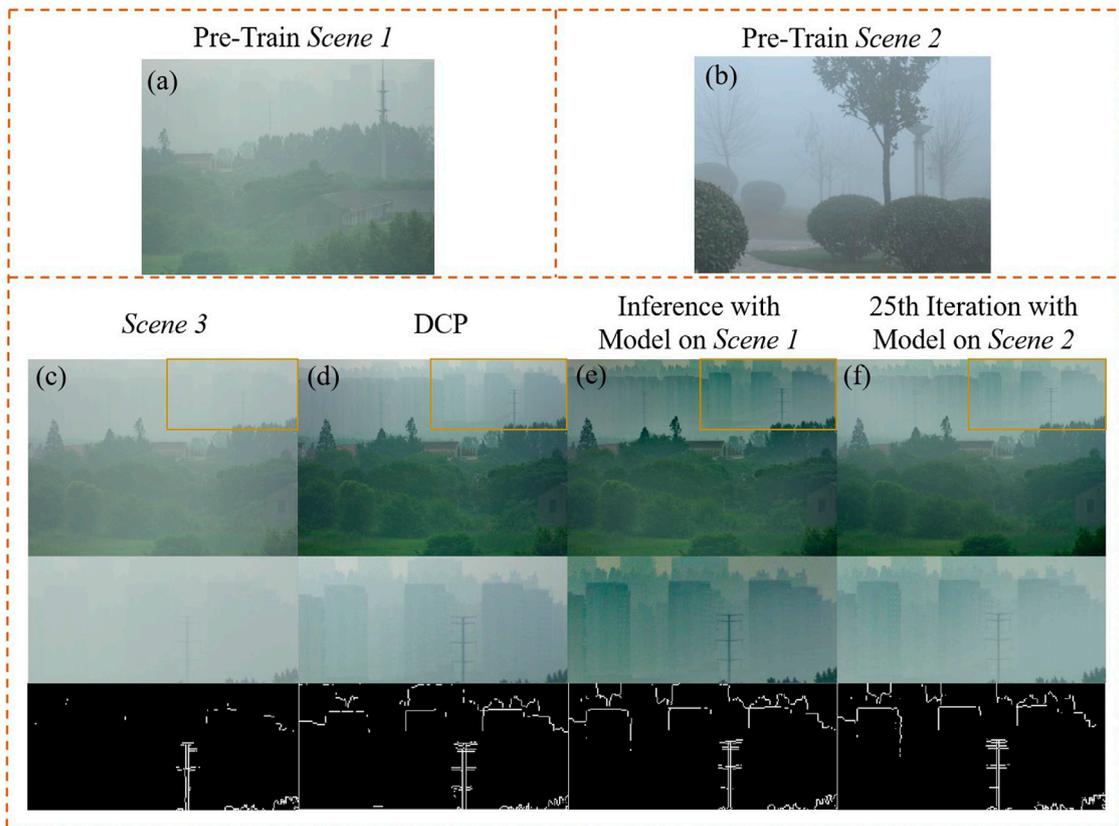


FIGURE 6 | Cross-scene dehazing capability comparison. **(A)** Pre-training scene 1. **(B)** Pre-training scene 2. **(C)** Scene 3 to be dehazed. **(D)** Dehazing results by using the DCP-based method. **(E)** Scene 3 dehazing results by using the scene 1 training model for direct inference. **(F)** haze removal results using the training model of scene 2 as a pre-training model and 25 gradients are back-propagation for updates. The distant scenes with severe image degradation are locally zoomed in and edge extraction is performed to compare the dehazing ability of different methods.

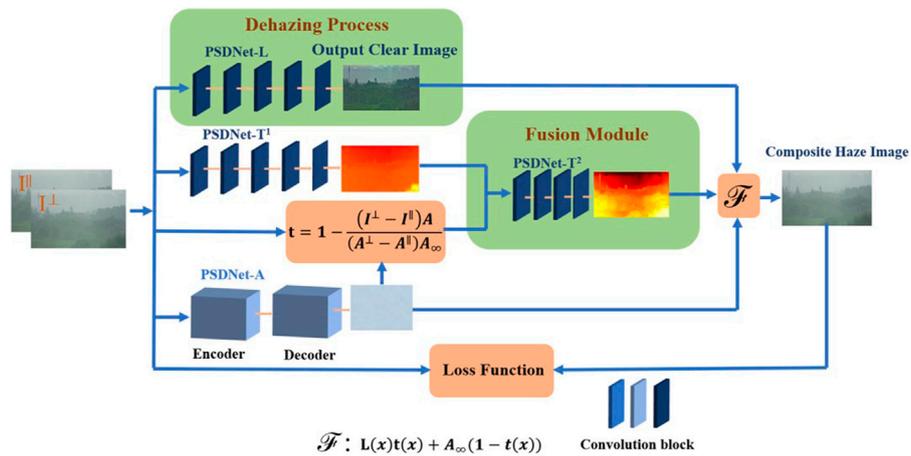


FIGURE 7 | Schematic diagram of PSDNet.

TABLE 3 | Comparison of calculating time using different methods.

Method	Computing platform	Preparation time(s)	Single inference/calculation time(s)
DCP	I7-9700	None	2.18
AODNet	I7-9700 and RTX titan	11.09	0.53
Ours	I7-9700 and RTX titan	13.53	0.34

Bold value indicate the best values.

superior dehazing outcomes than those obtained using the DCP method. In addition, when the pre-trained model is loaded, only the computational process of the *Dehazing Process* in **Figure 7** is required, so the running efficiency of the network can be greatly improved.

As mentioned above, supervised algorithms need a big quantity of data for training; aside from the collection of haze-free pairs that take a lot of time, the model training procedure takes a significant amount of computer resources and time. AODNet, for example, takes more than 36 h to train 50 epochs, but PSDNet takes less than 5 min to train 800 epochs in the same computational environment, and less than 10 s is needed to complete 25 epoch of training when the pre-trained model is loaded. PSDNet is unquestionably quicker than supervised algorithms.

As shown in **Table 3**, the time required by different algorithms for haze removal is compared. There is a preparation time since AODNet and PSDNet need to load the model to the GPU, but the model only has to be loaded once, and then the network can remove the haze of numerous pictures. In terms of time comparison, the single frame dehazing time of PSDNet only takes 0.34 s. The dehazing speed of PSDNet, which is significantly faster than AODNet and DCP-based methods, makes it possible to achieve quasi-real-time dehazing.

4 DISCUSSION

According to the experimental results, we have the following discussions.

- 1) PSDNet combined with the physical model can efficiently utilize the scene polarization information for accurate estimation of the transmission map and form a self-supervised closed loop. Therefore, haze-free images are not required as GT for constraint during all training processes, which reduces the dependence on data. Compared with the dehazing results of traditional methods using polarization and the dehazing results of supervised networks, PSDNet has better performance in enhancing scene details and color retention, and can almost achieve the enhancement from unseen to visible target in some scenes.
- 2) PSDNet is robust for different scenarios. Because the physical models included in the neural network are built based on actual haze scenes, PSDNet is effective at most scenes. And the network structure incorporated physical priors can help the models trained with different scenes to migrate or online learning. The training times can be

reduced to 1/32 of the original by loading the pre-trained model (from 800 epochs to 25 epochs, result as shown in **Figure 6F**). For similar scenes, the pre-trained model can be directly used to remove haze without retraining (result as shown in **Figure 6E**).

- 3) Because PSDNet does not require a large amount of data for training, this advantage not only drastically reduces the data acquisition time but also saves the time for model training. Compared to the supervised algorithm AODNet, which takes 36 h to train 50 epochs, PSDNet takes less than 5 min for 800 iterations, and the training time can be compressed to less than 10 s when loading the pre-trained model. When performing model inference for dehazing, PSDNet is three times faster than traditional methods based on DCP, and also faster than similar supervised class algorithms.

5 CONCLUSION

This paper proposes a method that combines the polarization difference of the scene with the neural network to achieve dehazing. Since the polarization prior can effectively guide and activate the extracted feature maps of neural networks, the proposed network does not need haze-free pairs as GT to constrain the training process. Only two frames of scene images with orthogonal polarization at any angle are required as input, then the self-supervision and global online optimization learning approach are used for haze removal. The airlight can be better estimated by the self-supervised closed-loop optimization process. Therefore, the proposed method has good results in preserving the color of the original image and enhancing the details compared to similar algorithms based on polarization or supervised learning-based. In actual dense haze scenes, almost invisible details of distant targets can be identified by using the proposed approach for dehazing. The training time and dehazing efficiency of the network have obvious advantages in the comparison of similar methods, and it is expected to achieve real-time haze removal. The proposal of this method promotes the development of the combination of deep learning and physical models in the field of anti-scattering.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

YS performed the numerical calculation and wrote the manuscript. EG and JH performed the data analysis and provided constructive discussions. EG, LB, and JH are the main supervisors, and they provided supervision and feedback and reviewed the research.

REFERENCES

- 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* (2016) 25:5187–98. doi:10.1109/tip.2016.2598681
- Zhang X, Dong H, Pan J, Zhu C, Tai Y, Wang C, et al. Learning to Restore Hazy Video: A New Real-World Dataset and a New Method. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2021 Jun 19–25 (2021). p. 9239–48. doi:10.1109/cvpr46437.2021.00912
- Dong H, Pan J, Xiang L, Hu Z, Zhang X, Wang F, et al. Multi-scale Boosted Dehazing Network with Dense Feature Fusion. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2020 Jun 19–25 (2020). p. 2157–67. doi:10.1109/cvpr42600.2020.00223
- He K, Sun J, Tang X. Single Image Haze Removal Using Dark Channel Prior. *IEEE Trans Pattern Anal Mach Intell* (2010) 33:2341–53. doi:10.1109/TPAMI.2010.168
- Bahat Y, Irani M. Blind Dehazing Using Internal Patch Recurrence. In: 2016 IEEE International Conference on Computational Photography (ICCP); 2016 May 13–15; Evanston, IL. IEEE (2016). p. 1–9. doi:10.1109/iccpht.2016.7492870
- Berman D, Treibitz T, Avidan S. Non-local Image Dehazing. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016 Jun 26–Jul 1; Las Vegas, NV (2016). p. 1674–82. doi:10.1109/cvpr.2016.185
- Schechner YY, Narasimhan SG, Nayar SK. Polarization-based Vision through Haze. *Appl Opt* (2003) 42:511–25. doi:10.1364/ao.42.000511
- Liang J, Ren L, Ju H, Zhang W, Qu E. Polarimetric Dehazing Method for Dense Haze Removal Based on Distribution Analysis of Angle of Polarization. *Opt Express* (2015) 23:26146–57. doi:10.1364/oe.23.026146
- Qu Y, Zou Z. Non-sky Polarization-Based Dehazing Algorithm for Non-specular Objects Using Polarization Difference and Global Scene Feature. *Opt Express* (2017) 25:25004–22. doi:10.1364/oe.25.025004
- Pang Y, Nie J, Xie J, Han J, Li X. Bidnet: Binocular Image Dehazing without Explicit Disparity Estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2020 Jun 19–25 (2020). p. 5931–40. doi:10.1109/cvpr42600.2020.00597
- Yang X, Xu Z, Luo J. Towards Perceptual Image Dehazing by Physics-Based Disentanglement and Adversarial Training. In: Proceedings of the AAAI Conference on Artificial Intelligence; 2018 Feb 2–7; New Orleans, LA, 32 (2018).
- Li B, Peng X, Wang Z, Xu J, Feng D. Aod-net: All-In-One Dehazing Network. In: Proceedings of the IEEE International Conference on Computer Vision; 2017 Oct 22–29; Venice (2017). p. 4770–8. doi:10.1109/iccv.2017.511
- Liu F, Cao L, Shao X, Han P, Bin X. Polarimetric Dehazing Utilizing Spatial Frequency Segregation of Images. *Appl Opt* (2015) 54:8116–22. doi:10.1364/ao.54.008116
- Fang S, Xia X, Huo X, Chen C. Image Dehazing Using Polarization Effects of Objects and Airlight. *Opt Express* (2014) 22:19523–37. doi:10.1364/oe.22.019523
- Van der Laan JD, Scrymgeour DA, Kemme SA, Dereniak EL. Detection Range Enhancement Using Circularly Polarized Light in Scattering Environments for Infrared Wavelengths. *Appl Opt* (2015) 54:2266–74. doi:10.1364/ao.54.002266
- Dudhane A, Singh Aulakh H, Murala S. Ri-gan: An End-To-End Network for Single Image Haze Removal. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops; 2019 Jun 16–20; Long Beach, CA (2019). doi:10.1109/cvprw.2019.00253
- Shen L, Zhao Y, Peng Q, Chan JCW, Kong SG. An Iterative Image Dehazing Method with Polarization. *IEEE Trans Multimedia* (2018) 21:1093–107. doi:10.1109/TMM.2018.2871955
- Shen L, Reda M, Zhao Y. Image-matching Enhancement Using a Polarized Intensity-Hue-Saturation Fusion Method. *Appl Opt* (2021) 60:3699–715. doi:10.1364/ao.419726
- Ren W, Liu S, Zhang H, Pan J, Cao X, Yang MH. Single Image Dehazing via Multi-Scale Convolutional Neural Networks. In: European Conference on Computer Vision; 2016 Jun 26–Jul 1; Las Vegas, NV. Springer (2016). p. 154–69. doi:10.1007/978-3-319-46475-6_10
- Nathan Silberman PK, Hoiem D, Fergus R. Indoor Segmentation and Support Inference from Rgb-d Images. In: European Conference on Computer Vision; 2012 Oct 7–13; Firenze (2012). doi:10.1007/978-3-642-33715-4_54
- Li B, Ren W, Fu D, Tao D, Feng D, Zeng W, et al. Benchmarking Single Image Dehazing and beyond. *IEEE Trans Image Process* (2018) 28:492–505. doi:10.1109/TIP.2018.2867951
- Chen L, Zhao J, Lin M, Yang G, He C. No-reference Perceptual Quality Assessment of Stereoscopic Images Based on Binocular Visual Characteristics. *Signal Processing: Image Commun* (2019) 76:1–10. doi:10.1016/j.image.2019.03.011

FUNDING

This research is supported by the National Natural Science Foundation of China (62031018, 61971227, 62101255); Jiangsu Provincial Key Research and Development Program (BE2018126).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Shi, Guo, Bai and Han. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.