Single-scale Residual Dense Dehazing Network

Nian Wang*, Aihua Li, Zhigao Cui, Yanzhao Su and Yunwei Lan
Xi’an Research Institute of High Technology, Xi’an, Shaanxi, China

*Corresponding author e-mail: 1150712112@wdd1949.114ic.com

Abstract. Recently, image dehazing algorithm has been widely used in the preprocessing of target tracking and pattern recognition. A large number of end-to-end convolutional networks have achieved good results in dehazing by image translation. In this paper, we use Residual Dense structure, which has been proved effective in high resolution reconstruction, to build feature extract block, and stack these block to form a single-scale dehazing network. In order to further enhance the performance, we convey the feature of shallow layer to deep layer by channel concatenation. The results show that our network has achieved good results in both the synthetic haze removal and the real haze removal.

Keywords: RDB, Single-scale, Channel Concatenation

1. Introduction

A range of advanced computer tasks rely on clear input images, such as outdoor monitoring and target tracking. However, due to the existence of particles such as haze and dust, many of the outdoor images we take suffer from a diverse degree of quality loss. To solve this problem, researchers have carried out a large number of experiments. For example, the traditional image dehazing algorithm depend on the atmospheric scattering model [1].

\[ I(x) = J(x) t(x) + A(1-t(x)) \]  

(1)

In Formula (1), \( I(x) \) denotes hazy image, \( J(x) \) denotes clear image, \( A \) and \( t(x) \) denote global atmospheric light and scene transmission respectively. If \( A \) and \( t(x) \) are known, a clear image can be obtained from the hazy image. However, it is an ill-posed problem. In order to solve this problem, researchers have proposed lots of assumptions to constrain the atmospheric scattering model [2, 3]. This method can restore the image quickly, but its quality is not high. For example, the dark channel prior [2] tends to underestimate the transmission map of the sky region, which leads to color saturation.

Recently, image dehazing based on data-driven [4] has become the mainstream. By building the convolutional neural network to learn transmission map, it gets higher quality of dehazed images. However, most of them still use artificial method to estimate the global atmospheric light, which leads to instability of dehazing results. So AOD-net [5] combines global atmospheric light and scene transmission as one parameter through linear transformation; DCPDN [6] embeds the atmospheric...
scattering model into Generative Adversarial Networks (GAN) to estimate scene transmission and global atmospheric light simultaneously. More recently, instead of using the atmospheric scattering model, most methods incline to directly establish the mapping relationship between hazy image and clear image [7, 8, 9]. This indicates image dehazing is similar to domain migration, super-resolution reconstruction, which is a problem to restore an image from another image with different feature distribution.

In this paper, we use Residual Dense structure [10], which combines Dense-net [11] with Res-net [12], to build an effective dehazing network on single-scale, and directly restore dehazed image from hazy image.

2. The Network Structure
It is generally agreed that the shallow convolutional layer can capture more high-frequency information, such as texture and edge, due to its small receptive field. While the deeper convolutional layer can capture more global information, such as light and color, due to its large perception field. For a deep network built on the same scale, channel fusion is needed to transmit the shallow information to the deep layer to prevent the loss of the detailed information. Based on this, we use the Residual Dense Block (RDB) shown in the Figure. 1.

![Figure 1. The structure of RDB](image)

We further build our Single-scale Residual Dense Dehazing Network (SRDDN) with the RDB block. As is shown in Figure 2, we stack ten RDB block and one convolution layer as a group to extract feature. In end of every group, we use element-wise sum to add its input. Our Group structure not only iterate the feature map of the previous layer, but also guarantee the gradient flow by the residual structure. Our SRDDN network extracts deep haze feature through three Group structure, and combine them effectively by channel concatenation. After that, we change the number of channels to three through two convolutional layers to obtain the preliminary result of dehazing. Finally, we use element-wise sum to add initial input hazy image, which ensure integrity of image structure information. All convolutional layers are designed to ensure that the resolution of the image keeps 16 x 16.
3. Experiment and Analysis

All the experiments are carried out on the computer with Intel(R) Core(TM) i7-7800X CPU @ 3.50 GHZ and NVIDIA GeForce RTX 2080 Ti.

3.1. Data Sets and Train Arrangement

We use synthetic haze data set (SOTS) to train our SRDDN network. In detail, we use ITS (indoor) and OTS (outdoor) sets to train and test our network respectively. For indoor images, we train 100 epochs in total and reduce the learning rate by half every 10 epochs. While for outdoor images, we train 10 epochs in total and the learning rate is reduced by half every 2 epochs. For validation, we choose the remain images of ITS and OTS to prove its dehazing performance on synthetic hazy image. What is more, two images are selected to further validate its dehazing ability of real haze.

Loss Fuction

To train our network, L1 loss and the perceptual loss [13] are employed. L1 loss retain the structural characteristics of the dehazed images. The perceptual loss use the intermediate feature map of the pre-training network as a reference, which makes the final results delight our senses. We use the VGG16 [14] pre-trained on ImageNet [15] and extract the features from each of the first three stages. The total loss can be expressed as

$$L_{total} = L_1 + \lambda L_p$$

Where $\lambda$ is weight coefficient, which is set to 0.05.

3.2. Comparison on SOTS

As is shown in Figure 3, the dehazed images of DCP (b), and DCPDN (d) suffers color distortion. For example, the desk is bright (see row 1) and the sky appears white halo (see row 4). AOD-net (c) and GFN (e) have residual haze, because the red wall is white (see row 1) and the portrait is fuzzy (see row 2). In contrast, the indoor results of grid-Net (f) and FFA (g) are more similar to the real image. However, both of them (c) (e) (f) (g) have residual haze in sky of the dehazed image (see row 4). Subjectively, our dehazing results are closest to the ground truth (compare h and i).
In order to furtherly compare the dehazing effect of above networks, the Peak-to-noise Ratio (PSNR) and Structural Similarity (SSIM) of images are calculated respectively. As shown in Table 1, our dehazing network is close to Grid-net on indoor set but better than it on outdoor set.

| SOTS | Metrics | DCP | AOD | DCPDN | GFN | Grid | FFA | Ours |
|------|---------|-----|-----|-------|-----|------|-----|------|
| indoor | PSNR | 19.85 | 19.69 | 23.17 | 22.42 | 32.16 | 36.36 | 32.14 |
|       | SSIM  | 0.875 | 0.847 | 0.885 | 0.914 | 0.983 | 0.993 | 0.987 |
| outdoor| PSNR | 20.44 | 22.08 | 22.17 | 21.41 | 30.76 | 33.57 | 32.87 |
|       | SSIM  | 0.924 | 0.898 | 0.867 | 0.858 | 0.981 | 0.984 | 0.982 |

3.3. Comparison on Real Haze Images

As shown in Figure 4, the dehazed images of DCP (b) are distorted and DCPDN (d) is brighter than normal image (see row 1). AOD-net (c), GFN (e), Grid-net (f) and FFA (g) have residual haze in thick haze area (see row 2). More importantly, the restored images (see row 4) are generally fuzzy and dark, and only the results of DCPDN (d) can see the trees beside the track (see row 4) since its brightness is overestimated (see row 1 and 3). In contrast, our results (h) are clearer than other algorithms (see row 2 and 4), and the clear image is not at the expense of false brightness.
Figure 4. Real haze images comparison.

(a) Hazy Image; (b) DCP; (c) AOD-net; (d) DCPDN; (e) GFN; (f) Grid-net; (g) FFA; (h) Ours;

4. Conclusion
We build a Single-scale Residual Dense Dehazing Network with the RDB block, named SRDDN. Our network stack the RDB block to ensure the feature flow from shallow layer to deep layer, and put all intermediate feature of groups by Channel-wise Sum. Experiment shows our network keep the structural information of hazy image with good dehazing effect in both the synthetic haze and the real haze sets. Especially on real haze removal, our results are quite better.

References
[1] Narasimhan S G, Nayar S K. Contrast restoration of weather degraded images. IEEE Transactions on Pattern Analysis Machine Intelligence, 2003, 25(6), pp. 713-724.
[2] He K, Sun J, Fellow, et al. Single Image Haze Removal Using Dark Channel Prior. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 33(12), pp. 2341-2353.
[3] Q. Zhu, J. Mai, and L. Shao. Single image dehazing using color attenuation prior. In Proc. British Machine Vision Conference (BMVC), 2014, 7(02), pp. 23578-23584.
[4] Bolun Cai, Xiangmin Xu, Kui Jia, et al. Dehazenet: An end-to-end system for single image haze removal. IEEE Transactions on Image Processing, 2016, 25(11), pp. 5187–5198.
[5] Boyi Li, Xiulian Peng, Zhangyang Wang, et al. Aod-net: All-in-one dehazing network. In IEEE International Conference on Computer Vision, 2017, pp. 4770–4778.
[6] He Zhang, Vishal M Patel. Densely connected pyramid dehazing network. In IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3194–3203.
[7] Liu X, Ma Y, Shi Z, et al. GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing. Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 7314-7323.
[8] Ren W, Ma L, Zhang J, et al. Gated fusion network for single image dehazing. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3253-3261.
[9] Qin X, Wang Z, Bai Y, et al. FFA-Net: Feature Fusion Attention Network for Single Image Dehazing. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7), pp. 11908-11915.
[10] Yulun Zhang, Yapeng Tian, Yu Kong, et al. Residual dense network for image super-resolution In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 2472–2481.
[11] Huang G, Liu Z, Laurens V D M, et al. Densely Connected Convolutional Networks. 2016, pp.
[12] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2016, pp. 770–778.

[13] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In European Conference on Computer Vision (ECCV), Springer, 2016, pp. 694–711.

[14] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[15] Olga Russakovsky, Jia Deng, Hao Su, et al. Imagenet large scale visual recognition challenge. International Journal of Computer Vision (IJCV), 2015, 115(3), pp. 211–252.