Recursive Deep Residual Learning for Single Image Dehazing

Yixin Du and Xin Li
West Virginia University
LCSEE, 395 Evansdale Drive, Morgantown, WV 26506-6070, U.S.A.
yidu@mix.wvu.edu Xin.Li@mail.wvu.edu

Abstract

There have been a flurry of works on deep learning based image dehazing in recent years. However, most of them have only used deep neural networks to estimate the transmission map (or its variant); while the module of generating dehazed image is still model-based. Inspired by the analogy between image dehazing and image denoising, we propose to reformulate dehazing as a problem of learning structural residue (instead of white Gaussian noise) and remove haze from a single image by a deep residue learning (DRL) network. Such novel reformulation enables us to directly estimate a nonlinear mapping from input hazy images to output dehazed ones (i.e., bypassing the unnecessary step of transmission map estimation).

The dehazing-denoising analogy also motivates us to leverage the strategy of iterative regularization from denoising to dehazing - i.e., we propose to recursively feed the dehazed image back to the input of DRL network. Such recursive extension can be interpreted as a nonlinear optimization of DRL whose convergence can be rigorously analyzed using fixed-point theory. We have conducted extensive experimental studies on both synthetic and real-world hazy image data. Our experimental results have verified the effectiveness of the proposed recursive DRL approach and shown that our technique outperforms other competing methods in terms of both subjective and objective visual qualities of dehazed images.

1. Introduction

Outdoor images often suffer from low contrast and degraded color fidelity due to haze, fog, and other atmospheric phenomena. These phenomena are caused by light scattering and absorption with dusk, smoke and dry particulates in the atmosphere. Haze attenuates the reflected signal of a scene in the physical world, introducing undesirable quality degradation to the image [2] as shown in Figure 1(a). How to restore the visual quality of hazy images is often called "image dehazing" or "haze removal", which has been widely studied by the computer vision community in recent years.

Haze removal from a single image is a challenging ill-posed problem for several reasons. First, as mentioned in previous works [17, 10, 2, 14, 15], since the transmission map depends on the unknown and spatially varying scene depth, it is difficult to obtain an accurate estimate of transmission map without the faithful depth information. Second, in addition to transmission map, global atmosphere light also interfere with the process of image formation (i.e., so-called airlight [17] depends on both global atmosphere light and transmission map). The variation of global atmosphere light (e.g., nighttime vs. daytime) makes single-image dehazing more difficult. Last but not least, most of previous works assumed a simplified additive image formation model in order to make the problem of haze removal more analytically tractable. How accurate this model matches the actual formation of a hazy image in the real world is largely unknown to the best of our knowledge.

Conventional wisdom for image dehazing heavily relies on the simplified observation model and the estimation of transmission map. Various strategies for estimating the unknown scene depth or transmission map have been developed in both model-based approaches (e.g., independent component analysis [8] and dark channel prior [10]) and more recently learning-based approaches (e.g., MSCNN [20] and DehazeNet [3]). Inspired by the analogy between image dehazing and image denoising, we propose to reformulate dehazing as a problem of learning structural residue (instead of white Gaussian noise) and remove haze from an image by a deep residue learning (DRL) network. Such novel reformulation enables us to directly estimate a nonlinear mapping from input hazy images to output dehazed ones (i.e., bypassing the unnecessary step of transmission map estimation).

The dehazing-denoising analogy also motivates us to leverage the strategy of iterative regularization [19] from denoising to dehazing - i.e., we propose to recursively feed the dehazed image back to the input of DRL network. Such recursive extension can be interpreted as a nonlinear opti-

(a) Input (b) DCP [10] (c) NLD [2]

(d) MSCNN [20] (e) AOD [14] (f) DRL

Figure 1: Visual quality comparison between our proposed DRL and other state-of-the-art methods.

mization of DRL whose convergence can be rigorously analyzed using fixed-point theory [9] (haze-free images are the fixed points of learned nonlinear mappings by DRL network). We have conducted extensive experimental studies on both synthetic and real-world hazy image data. Our experimental results have verified the effectiveness of the proposed recursive DRL and shown that our approach outperforms other competing methods in terms of both subjective and objective visual qualities of dehazed images. Moreover, previous learning-based dehazing approaches have used a large number of training data (e.g., AOD [14] used 27,256 images), whereas our DRL achieves good performance with only 1,200 training images thanks to the patch extracting strategy (therefore requiring much fewer computational resources).

2. Related Work

Model-based approaches toward single image haze removal are often based on some haze-relevant priors. In [10], a Dark Channel Prior (DCP) was proposed based on the observations that in a haze-free image, any local patch would contain some pixels of low intensity values in at least one color channel. In [18], a factorial Markov Random Field model was adopted to jointly estimate the scene albedo and depth. In [15], the inherent boundary constraint of the transmission function was exploited during the estimation, which has been called contextual regularization. In nonlocal image dehazing [2], distance map and haze-free images are jointly estimated from haze-lines characterizing the linear-spreading structure of pixels within a given cluster in the RGB space. All those dehazing methods estimate the scene depth or transmission map to facilitate model-based image restoration. Due to the effectiveness of haze-relevant priors, most of the state-of-the-art model-based approaches show strong potentials in the heavy haze area of an image, at the cost of longer processing time.

Inspired by the success of deep learning in various low-level vision tasks such as image super-resolution [6, 7] and image denoising [25, 26], learning-based approaches have also been proposed for singe image dehazing in recent years [20, 3, 14]. The common foundation behind those works is the creation of training data - a large number of synthetic hazy images based on ground truth depth information. In [20], a multi-scale convolutional neural network (MSCNN) was proposed to estimate transmission map from an input of hazy image. In [3], authors developed an end-to-end dehazing network called DehazeNet in which network layers are specially tailored to fit assumptions/priors in the scenario of image dehazing. Most recently, the so-called All-in-One Dehazing (AOD)[14] further develops this line of idea by unifying the estimation of transmission map and global atmosphere light into one module called K-estimation.

3. Image Dehazing via Deep Residue Learning

The motivation behind this work is largely two-fold. On one hand, we want to pursue a direct approach of learning a nonlinear mapping from the space of hazy images to that of dehazed ones (i.e., to bypass the step of estimating transmission map) that is less dependent on the approximated assumption made about the model of hazy image formation. The feasibility of such a direct approach is supported by successful applications of deep learning into a variety of low-level vision tasks from denoising to super-resolution. On the other hand, there have been a flurry of works on deep residue learning in recent years (e.g., [11]) which have shown learning a residue representation is more effective than learning an image one. Such idea inspires us to reformulate image dehazing as a variant of image denoising problem as we will elaborate next.

3.1. Analogy between Image Dehazing and Denoising

To facilitate our discussion, we start from the simplified observation model for a hazy image:

$$\mathbf{I}(x) = \mathbf{J}(x)t(x) + \mathbf{A}(1 - t(x)), \tag{1}$$

t where $\mathbf{I}$ is the observed hazy image, $\mathbf{J}$ is the scene radiance (target of restoration), $\mathbf{A}$ is the global atmospheric light, and $t(x)$ is the medium transmission map characterizing the amount of attenuation. On the right-hand side of Eq. (1), $\mathbf{J}(x)t(x)$ represents the term of *direct attenuation* [23, 10, 16], and $\mathbf{A}(1-t(x))$ is called *airlight* [23, 10, 16]. Direct attenuation accounts for decay of scene radiance through the transmission medium; and airlight is attributed to atmospheric scattering [10] (together they create the hazy effect). The transmission map $t(x)$ is related to the scene depth by:

$$t(x) = e^{-\beta d(x)}, \tag{2}$$

where β is the scattering coefficient of the atmosphere and $d(x)$ is the depth of the scene at location x. It is easy to see that the amount of attenuation depends on the distance from the scene point to the camera [10] - the larger the distance, the stronger the attenuation. Additionally, the global atmospheric light would interfere the process of image degradation, which make image dehazing more challenging in night-time than in day-time.

In previous approaches, Estimating $\mathbf{A}$ and $t(x)$ are necessary steps for image dehazing; in this work, we propose to bypass them and take a direct approach of learning a nonlinear mapping from $\mathbf{I}(x)$ to $\mathbf{J}(x)$. Our direct approach is best illustrated based on an analogy between image dehazing and image denoising. If we reformulate Eq. (1) as follows:

$$\begin{aligned} \mathbf{I}(x) &= \mathbf{J}(x) + (\mathbf{A} - \mathbf{J}(x))(1 - t(x)) \\ &= \mathbf{J}(x) + \mathbf{r}(x) \end{aligned} \tag{3}$$

where $\mathbf{r}(x) = (\mathbf{A} - \mathbf{J}(x))(1 - t(x))$ can be interpreted as an error term characterizing the nonlinear signal-dependent degradation associated with the hazy effect. Such reformulation enables us to connect image dehazing with the widely-studied problem of image denoising. Recall in image denoising, we have

$$\mathbf{I}(x) = \mathbf{J}(x) + \mathbf{w}(x) \tag{4}$$

where $\mathbf{I}(x), \mathbf{J}(x)$ denote noisy and clean images respectively, the additive noise term $\mathbf{w}(x) \sim N(0, \sigma_w^2)$ is assumed to be white Gaussian. By comparing Eq. (4) and Eq. (3), we observe the apparent similarity except the difference on the error term. Instead of dealing with additive white Gaussian noise, we work with $\mathbf{r}(x)$ - signal-dependent noise that has a nonlinear dependency with both attenuation process $\mathbf{A}$, $t(x)$ and the target of estimation $\mathbf{J}(x)$. Therefore, we can imagine a deep learning approach toward image denoising can be leveraged to the scenario of image dehazing.

3.2. Deep Residue Learning (DRL) Network for Image Dehazing

In early works of image denoising via deep neural networks, a direct mapping from noise image $\mathbf{I}(x)$ to clean image $\mathbf{J}(x)$ is learned from the training data. Such direct learning strategy suffers from the notorious vanishing gradient problem as described in [1] due to the long-term dependency between $\mathbf{I}(x)$ and $\mathbf{J}(x)$; the vanishing gradient problem becomes even more severe as the network depth increases. A better strategy is to learn the residue representation $\mathbf{w}(x) = \mathbf{I}(x) - \mathbf{J}(x)$ instead; such idea of deep residue learning (DRL) was first proposed for image recognition [11] and then adopted for image super-resolution in [12, 13] and image denoising [26].

Based on the analogy between dehazing and denoising, we propose to take a deep residue learning approach toward image dehazing here. More specifically, we aim at learning a nonlinear mapping Ω from $\mathbf{I}(x)$ to $\mathbf{r}(x) = \mathbf{I}(x) - \mathbf{J}(x)$. Then the dehazed image can be recovered via $\mathbf{J}(x) = \mathbf{I}(x) - \mathbf{r}(x)$ where $\mathbf{r}(x) = \Omega(\mathbf{I}(x))$. Similar to [26], we have adopted the following loss function to learn the parameters Θ in the proposed DRL network:

$$\mathcal{L}(\Theta) = \frac{1}{2N} \sum_{k=1}^{N} \|\Omega(\mathbf{I}_k(x); \Theta) - (\mathbf{I}_k(x) - \mathbf{J}_k(x))\|_F^2 \tag{5}$$

where $\{(\mathbf{I}_k(x), \mathbf{J}_k(x))\}_{k=1}^{N}$ represents N hazy and original training image (patch) pairs.

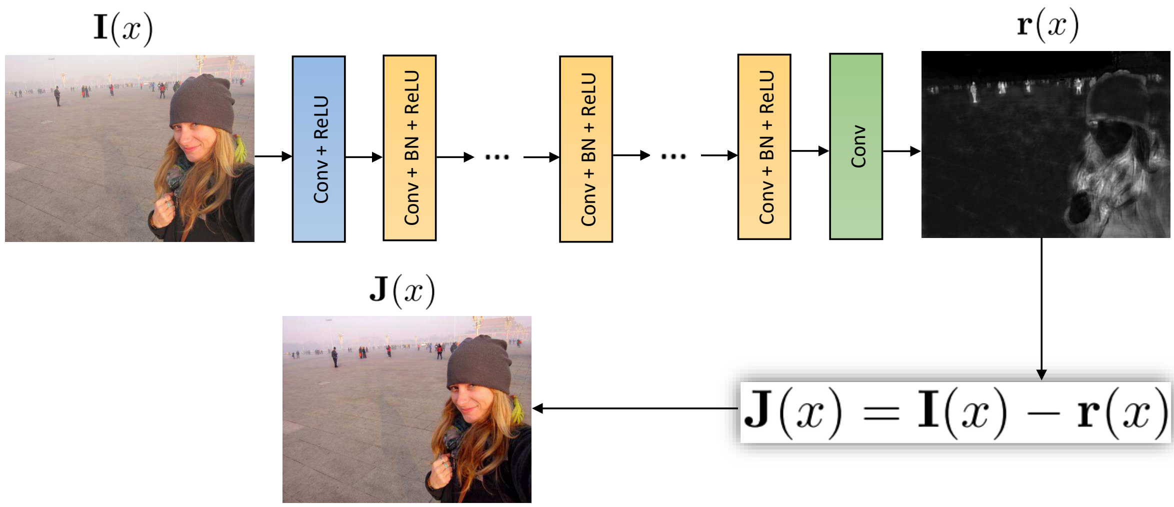

Figure 2: The architecture of the proposed DRL network.

The network architecture is shown in Figure 2. It contains three types of neural network layers: namely Convolutional (Conv), Rectified Linear Unit (ReLU), and Batch Normalization (BN). The first Conv layer consists of 32 feature maps generated by 32 filters sized by $3 \times 3 \times 3$ (since an input image has three color channels) and is followed by the Relu layer which performs $max(0, x)$ operation adding non-linearity to the model. Starting from the second Conv layer (as highlighted by color brown), we use 32 filters of size $3 \times 3 \times 32$ for each Conv layer, followed by BN and ReLU layers. The Conv + BN + ReLU concatenation is repeated for 15 times. Those network parameters such as filter size and network depth were empirically tuned to be nearly optimal. The last Conv layer (as highlighted by color green) includes 3 filters of size $3 \times 3 \times 32$ to calculate the loss between the outputs and training labels.

When compared with previous deep learning based approaches toward image dehazing (e.g., DehazeNet [3] and All-in-One Dehazing [14]), ours has several advantages. First, based on the vanishing gradient argument [1], the less correlated between the input and the output, the better the training outcome. Therefore, learning a residue representation such as $\mathbf{r}(x)$ is more effective than learning transmission map $t(x)$ or its variation (still highly correlated with the hazy input image). Second, since deep residue learning has bypassed the step of estimating transmission map, our approach actually does not rely on the validity of the assumed observation model in Eq. (1). In other words, even with an arbitrary nonlinear mapping $\mathbf{I}(x) = \Phi(\mathbf{J}(x))$ where Φ denotes the forward operator, our approach is still possible to learn an approximation of inverse operator Φ^{-1} by $\mathbf{1} - \Omega$ where $\mathbf{1}$ denotes the identity operator. Last but not the least, our direct DRL approach admits further optimization using the idea of iterative regularization [19], which we will explain next.

3.3. Image Dehazing Optimization via Fixed-point Iteration

The analogy between dehazing and denoising also leads to a nonlinear optimization of the proposed DRL-based dehazing via iterative regularization [19, 4]. In image denoising, suppose a regularized estimate of clean image from noisy observation $\mathbf{I}(x)$ is given by a nonlinear mapping $\Phi^{-1}(\mathbf{I}(x))$ and the error is denoted by $\mathbf{e}(x) = \mathbf{I}(x) - \Phi^{-1}(\mathbf{I}(x))$. If $\mathbf{e}(x)$ is already white Gaussian, then we are already done; otherwise (i.e., in the presence of leftover image structures), a simple strategy of improving upon this regularized estimate is to feed the denoised image $\Phi^{-1}(\mathbf{I}(x))$ back to the denoising algorithm and see if the new error is closer to white Gaussian. Apparently, such process can be recursively applied until the convergence (i.e., white Gaussian is the fixed point of denoising operator Ω), which has been called iterative regularization in the literature.

Similarly, we can recursively feed a dehazed image back to the input of the proposed DRL network as a strategy of nonlinear optimization. If the haze-free image is the fixed-point of our DRL network, the learned residue image $\Omega(\mathbf{I}(x))$ should asymptotically goes to zero. We have indeed empirically found that the learned nonlinear mapping $\mathbf{1} - \Omega$ satisfies the condition of a *nonexpansive* map (at least in the first few iterations): let K be a subset of a Banach space, a transformation $F : K \rightarrow K$ is said to be nonexpansive if for arbitrary $x, y \in K$, we have $||Fx - Fy|| < ||x - y||$. The existence of fixed point for nonexpansive mapping is given by the following generalization of Browder's theorem (Theorem 1 in [9])

Theorem 1. Let K be a nonempty, closed, convex and bounded subset of a uniformly convex Banach space X, and let $F : K \to K$ be asymptotically nonexpansive, then F has a fixed point.

The above theorem guarantees the theoretical existence of a fixed point for any nonexpansive map. However, we note that in practice it is often convenient and desirable to stop the recursion after a fixed number of iterations (or using so-called discrepancy principle as a stopping criterion [19]). Figure 3 illustrates the output of DRL network after the first three iterations. It can be observed that each iteration removes a portion of the haze. Depending on the amount of haze in the image, typically the scene radiance recovered after three to five iterations looks the most visually appealing.

4. Experimental Results and Comparisons

We have used the NYU-Depth V2 dataset [22] to create synthetic training images. NYU-Depth V2 dataset consists of 1,449 densely labeled indoor color images with ground truth depth information. The raw depth map has been projected and colorized [22] to fill in missing depth labels. Both the color and depth data are of the size 640 $\times$ 480. We pick 1,200 out of 1,449 images to generate training patches, and take the remaining 249 images as **ImageSet A**; we pick another 21 images from the Middlebury Stereo Datasets [21] as **ImageSet B**. For each image in the training set, we extract 40 $\times$ 40 patches with stride number being 30; there are 360,064 training patches generated in total.

To simulate synthetic hazy images, the following parameters are used in our experiment. We randomly select $\beta \in \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5\}$ since any value of β beyond this range could lead to unrealistic haze (too thin or too heavy) and noise amplification [20]. For each of the RGB channel, atmospheric light A is chosen uniformly within the range of $[0.6, 1.0]$. Training labels are generated using Eq. (3). During the training process, the weights of each convolution layers are randomly initialized by Gaussian variables. The number of epocs is set to 100; the learning rates for the first 60 and the remaining 40 epocs are set to 0.001 and 0.0001 respectively. We have selected Stochastic Gradient Descent (SGD) as the solver with a momentum parameter of 0.9. The network is trained on a PC with an Intel i7-4790k processor and a Nvidia GeForce Titan GPU leading to the total training time of about 15 hours.

4.1. Experimental Results on Synthetic Data and Real-world Images

In order to show the effectiveness of the proposed DRL, we compare our network with several state-of-the-art dehazing methods: Dark Channel Prior (**DCP**) [10], Boundary Constrained Context Regularization (**BCCR**) [15], Visual Artifact Suppression via Gradient Residual Minimization (**VASGRM**) [5], Non-Local Image Dehazing (**NLD**) [2], **MSCNN** [20], **DehazeNet** [3], and **AOD** [14]. The first four methods, including DCP, BCCR, VASGRM and NLD, are traditional model-based approaches, and the last three are learning-based approaches. Two objective image quality metrics are used in our comparison: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM) [24]. Table 1 and 2 shows the average PSNR and SSIM comparison results on ImageSet A and B respectively. Overall, we have observed that learning-based approaches, such as MSCNN, DehazeNet, and AOD, produce slightly better PSNR/SSIM results than model-based approaches; while our DRL outperforms all others by a large margin. We believe that dramatic performance improvement is jointly contributed by deep residue learning and fixed-point iterations.

Figure 4 shows the performance of our method compared against AOD [14] on ImageSet A and B. Since the NYU-Depth V2 dataset has been collected under indoor environment, most of the images have busy background with furniture and objects. Our method can handle such busy background very well - the scene radiance recovered maintains good sharpness with haze residual being close to zero. Our dehazed results are convincingly better than those produced by AOD on this data set. The main challenge of dehazing on ImageSet B is to recover rich details around sharp edges and vivid colors in those Middlebury images. Figure 4 shows our method can more faithfully maintain important image structure such as corners and edges and better recover the color fidelity than previous work of AOD.

We have also compared our method with seven competing dehazing approaches on real-world hazy images as shown in Figure 5 (a). This set of images - containing a large variation of scene content such as portrait, landscape and architecture - come from our own collection of real-world images that have been used in previous studies. These images contain both heavy and thin haze, shallow and large depth field, coarse and fine details, which reflect the diverse challenges in the real world. As we can see from Figure 5, our proposed technique has achieved at least comparable (and often superior) visual quality to other competing approaches.

4.2. Running Time Comparisons

Table 3 shows the comparison of running time (in seconds) among eight competing dehazing techniques. The results are obtained by running each method on ImageSet A which includes 249 images and taking their average. We have found that our method is relatively fast with 0.89 second per image (it is only marginally slower than AOD [14]). We believe this difference is mainly because of the platform, since Pycaffe is better optimized than Matlab in terms of implementing deep convolutional neural networks.

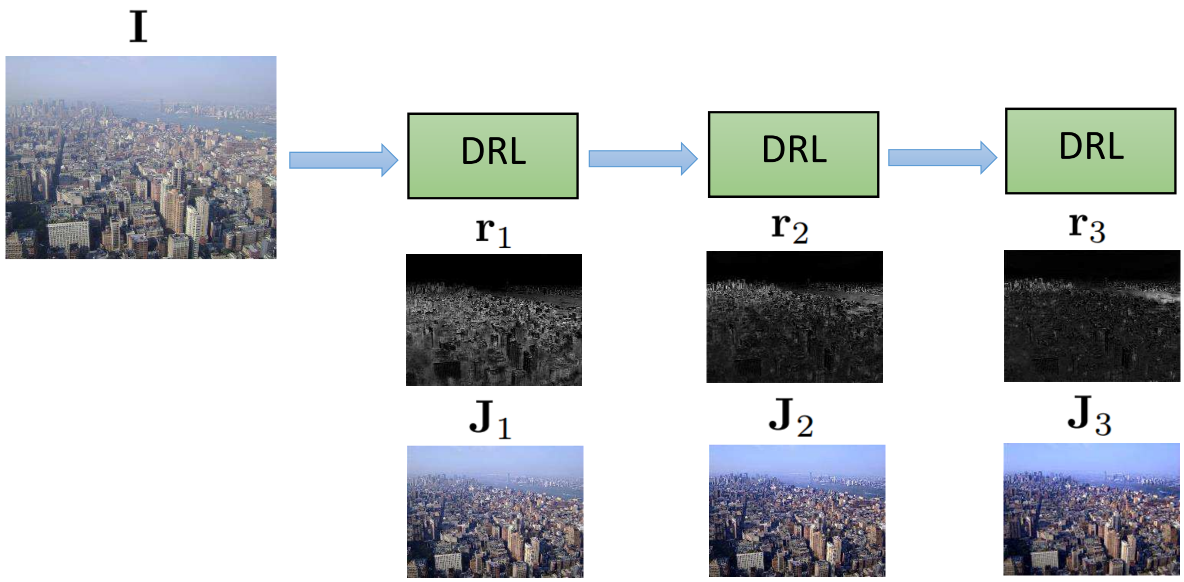

Figure 3: Recursive deep residue learning for haze removal: $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{r}_3$ are residuals recovered in the first three iterations; $\mathbf{J}_1$, $\mathbf{J}_2$ and $\mathbf{J}_3$ are recovered scene radiances.

Table 1: Average PSNR and SSIM results on ImageSet A.

Metrics	DCP [10]	BCCR [15]	VASGRM [5]	NLD [2]	MSCNN [20]	DehazeNet [3]	AOD [14]	**DRL**
PSNR (dB)	17.94	14.77	16.25	15.97	19.23	15.35	18.36	**21.7**
SSIM	0.86	0.81	0.83	0.77	0.86	0.76	0.85	**0.92**

Table 3: Comparison of average running time in seconds.

Method	Time	Platform
DCP [10]	8.3	Matlab
BCCR [15]	1.6	Matlab
VASGRM [5]	18.1	Matlab
NLD [2]	2.4	Matlab
MSCNN [20]	0.98	Matlab
DehazeNet [3]	1.4	Pycaffe
AOD [14]	0.53	Pycaffe
DRL	**0.87**	Matlab

5. Conclusion

This paper tackles image dehazing problem by taking a deep residue learning approach. By reformulating dehazing as a variation of denoising (with structured noise component), we propose to directly learn a nonlinear mapping from the space of hazy images to that of dehazed ones (i.e., bypassing the step of transmission map estimation). The analogy between dehazing and denoising also inspires us to leverage the idea of iterative regularization from denoising to dehazing leading to a nonlinear optimization by fixed-point iteration. Excellent experimental results have been reported to support the superiority of this work in all aspects (higher visual quality, smaller training set and faster running speed).

References

[1] Y. Bengio, P. Simard, and P. Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157–166, 1994. 3, 4

[2] D. Berman, S. Avidan, et al. Non-local image dehazing. In *Proceedings of the IEEE conference on CVPR*, pages 1674–1682, 2016. 1, 2, 5, 6, 7

[3] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11):5187–5198, 2016. 1, 2, 4, 5, 6, 7

[4] M. R. Charest and P. Milanfar. On iterative regularization and its application. *IEEE Transactions on CSVT*, 18(3):406–411, 2008. 4

[5] C. Chen, M. N. Do, and J. Wang. Robust image and video dehazing with visual artifact suppression via gradient residual minimization. In *ECCV*, pages 576–591. Springer, 2016. 5, 6, 7

Table 2: Average PSNR and SSIM results on ImageSet B.

Metrics	DCP [10]	BCCR [15]	VASGRM [5]	NLD [2]	MSCNN [20]	DehazeNet [3]	AOD [14]	**DRL**
PSNR (dB)	15.82	14.49	15.57	16.33	18.16	19.43	17.51	**21.41**
SSIM	0.81	0.77	0.81	0.79	0.84	0.85	0.85	**0.86**

(a) Inputs (b) DCP (c) BCCR (d) VASGRM (e) NLD (f) MSCNN (g) DehazeNet (h) AOD (i) DRL

Figure 5: Dehazing results on real-world images.

[6] C. Dong, C. C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In *ECCV*, pages 184–199. Springer, 2014. 2

[7] C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. *IEEE Transactions on PAMI*, 38(2):295–307, 2016. 2

[8] R. Fattal. Single image dehazing. *ACM transactions on graphics (TOG)*, 27(3):72, 2008. 1

[9] K. Goebel and W. Kirk. A fixed point theorem for asymptotically nonexpansive mappings. *Proceedings of the American Mathematical Society*, 35(1):171–174, 1972. 2, 4

[10] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. *IEEE transactions on PAMI*, 33(12):2341–2353, 2011. 1, 2, 3, 5, 6, 7

[11] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on CVPR*, pages 770–778, 2016. 3

[12] J. Kim, J. Kwon Lee, and K. Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE Conference on CVPR*, pages 1646–1654, 2016. 3

[13] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang. Deep laplacian pyramid networks for fast and accurate super-resolution. *Proceedings of the IEEE conference on CVPR*, 2017. 3

[14] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng. Aod-net: All-in-one dehazing network. In *Proceedings of the IEEE Conference on CVPR*, pages 4770–4778, 2017. 1, 2, 4, 5, 6, 7, 8

[15] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan. Efficient image dehazing with boundary constraint and contextual regularization. In *Proceedings of the IEEE ICCV*, pages 617–624, 2013. 1, 2, 5, 6, 7

[16] W. E. K. Middleton. Vision through the atmosphere. In *Geophysik II/Geophysics II*, pages 254–287. Springer, 1957. 3

[17] S. G. Narasimhan and S. K. Nayar. Vision and the atmosphere. *International Journal of Computer Vision*, 48(3):233–254, 2002. 1

[18] K. Nishino, L. Kratz, and S. Lombardi. Bayesian defogging. *International Journal of Computer Vision*, 98(3):263–278, 2012. 2

[19] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. *Multiscale Modeling & Simulation*, 4(2):460–489, 2005. 1, 4, 5

[20] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang. Single image dehazing via multi-scale convolutional neural networks. In *ECCV*, pages 154–169. Springer, 2016. 1, 2, 5, 6, 7

[21] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages I–I. IEEE, 2003. 5

[22] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus. Indoor segmentation and support inference from rgbd images. *ECCV*, pages 746–760, 2012. 5

[23] R. T. Tan. Visibility in bad weather from a single image. In *Proc. of CVPR*, pages 1–8. IEEE, 2008. 3

[24] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004. 5

[25] J. Xie, L. Xu, and E. Chen. Image denoising and inpainting with deep neural networks. In *NIPS*, pages 341–349, 2012. 2

[26] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 2017. 2, 3

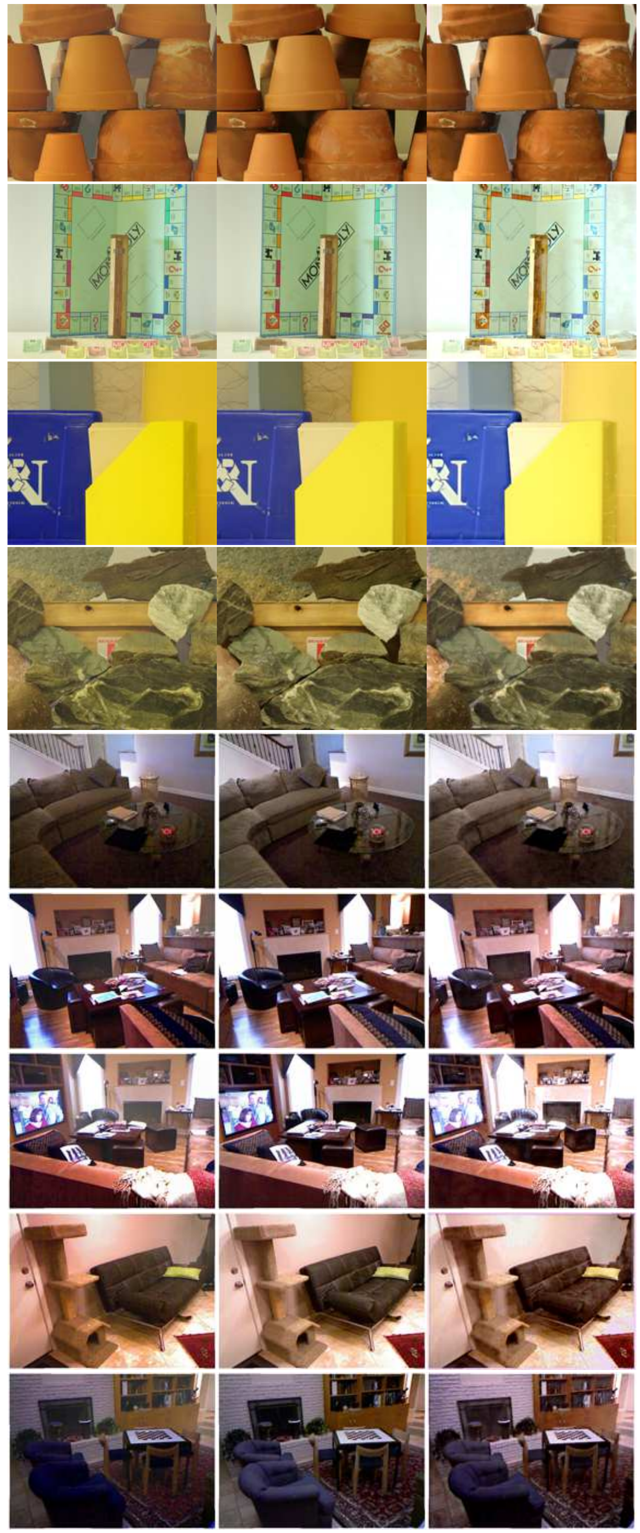

Figure 4: Dehazing results on ImageSet A and B. Left: AOD results [14]. Middle: original images. Right: outputs of DRL.