https://doi.org/10.2352/issn.2694-118X.2020.LIM-11

Single Image Dehazing by Predicting Atmospheric Scattering Parameters

Simone Bianco, Luigi Celona*, Flavio Piccoli
Department of Informatics, Systems and Communication, University of Milano - Bicocca, viale Sarca, 336 Milano, Italy
{***first_name.last_name@unimib.it***}

Abstract

In this work we propose a method for single image dehazing that exploits a physical model to recover the haze-free image by estimating the atmospheric scattering parameters. Cycle consistency is used to further improve the reconstruction quality of local structures and objects in the scene as well. Experimental results on four real and synthetic hazy image datasets show the effectiveness of the proposed method in terms of two commonly used full-reference image quality metrics.

Introduction

With the increasing demand for self-driving cars and high-fidelity cameras, the task of removing haze to improve the overall visibility of the subjects depicted in the scene has gained a lot of attention in the field of computer vision. Haze is the effect of several atmospheric events and consists of small-diameter particles having a density similar to that of the air, causing their floating. The presence of these particles causes an irregular deviation and the attenuation of the photons, resulting in an odd alteration of the light flow. Images acquired in such conditions present a degradation in the structures of the objects depicted and a strong decrease in contrast.

Two predominant strategies are usually used to restore the hazy image: the first one is to fit physical models that describe the hazing effects, while the second one is to perform an agnostic image-to-image restoration. Both strategies led to good results, indicating that even without directly specifying a physical model, it is possible to remove the hazing effect. Approaches using a physical model, however, not only restore the input image but can infer relevant factors involved in the process such as the current haze transmittance in the visible range. These side information can play a crucial role in determining what the network is inferring.

A valid physical model that adequately describes the hazing effect has been proposed by Mc Cartney *et al.* [20]. This model, namely the *atmospheric scattering model*, combines the scene radiance of the haze-free input with the global atmospheric illumination through the estimation of the haze transmittance as follows:

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

Specifically, the hazy image $\mathbf{I}(\mathrm{x})$ is composed by a weighted sum of the haze-free image $\mathbf{J}(\mathrm{x})$ and the ambient light $\mathbf{A}(\mathrm{x})$. The transmission map $t(\mathrm{x})$ defines the combination weights and is defined as:

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

where β is the scattering coefficient of the atmosphere, and $d(\mathrm{x})$ is the distance between the object in the scene projected at the spatial location x and the camera. Ambient light $\mathbf{A}(\mathrm{x})$ has three channels while the transmission map $t(\mathrm{x})$, which acts as a gating function, has only one channel. Several methods [6, 7, 12] consider the ambient light as uniform. In this case, $\mathbf{A}(\mathrm{x})$ has the same value for each location of the scene. He *et al.* [14] makes use of the *dark channel prior* to improve the contrast. The hypothesis behind this prior states that in non-sky regions of outdoor images there are some spots (the *dark pixels*) with very low intensity. The value of the dark pixel is directly affected by the airlight, therefore it directly provides an accurate estimation of the haze transmission. Beside physical-model-based methods, there are several other algorithms that learn a direct mapping between hazy image and their haze-free counterpart. These methods are usually trained using a perceptually-relevant metric such as [13], which uses Retinex, [24] which uses SSIM and finally HR-Dehazer [9], which uses a multiscale version of perceptual loss, all in a supervised fashion.

The main contribution of this paper is twofold:

- A method treating the dehazing problem as an image decomposition problem, where the hazy image is separated into its underling haze-free image and the obscuring foggy layer by predicting the atmospheric scattering parameters, namely the transmission matrix and the atmospheric illumination map;
- The estimation of a non-uniform atmospheric illumination map which considers local atmospheric color changes.

Proposed Method

In this work we treat the dehazing problem as an image decomposition problem, where we attempt to separate the hazy image into its underling haze-free layer and the obscuring foggy layer, respectively. We achieve this goal by predicting the atmospheric scattering parameters, namely the transmission matrix and the *non-uniform* atmospheric illumination map.

Figure 1 summarizes the proposed method for single image dehazing. Given a hazy color image $\mathbf{I}(x)$, first its RGB channels are projected into the monomial basis through a polynomial transformation, the resulting image is then fed into an encoder-decoder network which predicts the unknown parameters of the atmospheric scattering model, namely the transmission matrix $t(\mathrm{x})$ and the non-uniform atmospheric illumination map $\mathbf{A}(\mathrm{x})$. Finally, the haze-free image $\hat{\mathbf{J}}(\mathrm{x})$ is obtained by computing the inverse function of the atmospheric scattering model. Inspired by [10], given an input image, the (x_R, x_G, x_B) components of each pixel x are projected into the monomial basis $(1, x_R, x_G, x_B, x_R x_G, x_R x_B, x_G x_B, x_R^2, x_G^2, x_B^2, \ldots, x_R^D, x_G^D, x_B^D)$ for a set of degree D. The result of this operation, named "polynomial expansion", is an image of $\binom{D+3}{D}$ channels. We use the same encoder-decoder network defined in [9] for simultaneously predicting the transmission matrix and the non-uniform atmospheric illumination map. The first network output is a single-channel

* Corresponding author

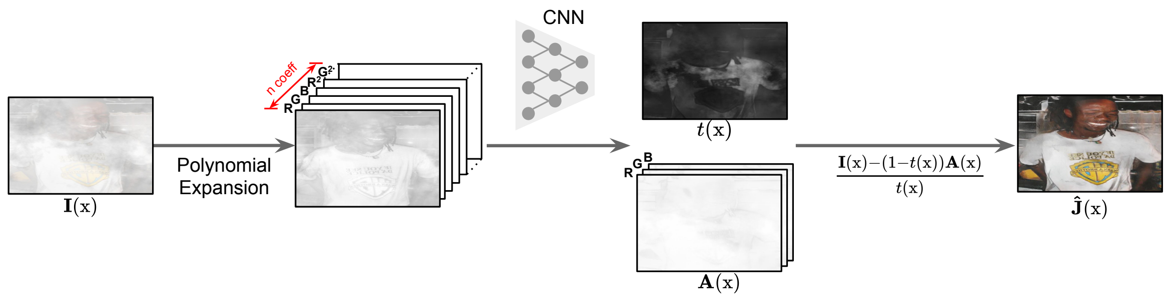

Figure 1: **Pipeline of the dehazing process.** The input hazy image $\mathbf{I}(\mathrm{x})$ is expanded by using a polynomial transformation, then the resulting image is fed into an encoder-decoder network for the estimation of the transmission $t(\mathrm{x})$ and the atmospheric light $\mathbf{A}(\mathrm{x})$ maps. Finally, the inverse equation of the *atmospheric scattering model* is computed for obtaining the haze-free image $\hat{\mathbf{J}}(\mathrm{x})$.

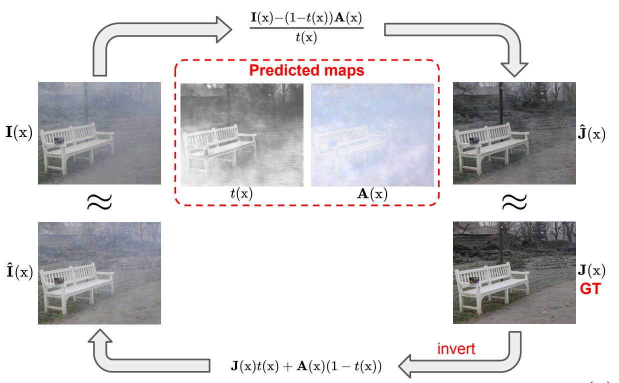

Figure 2: **Cycle-consistency.** Given the input hazy image $\mathbf{I}(\mathrm{x})$, the transmission $t(\mathrm{x})$ and the light $\mathbf{A}(\mathrm{x})$ maps are predicted to obtain the haze-free image $\hat{\mathbf{J}}(\mathrm{x})$ that should be equal to the ground-truth one $\mathbf{J}(\mathrm{x})$. Conversely, given the previously predicted maps, a hazy image $\hat{\mathbf{I}}(\mathrm{x})$ equal to the input image $\mathbf{I}(\mathrm{x})$ should be obtained by inverting the transformation.

image, while the second one is a three-channels image. In order to obtain the enhanced haze-free image, we apply the inverse atmospheric scattering model equation:

$$\hat{\mathbf{J}}(\mathrm{x}) = \frac{\mathbf{I}(\mathrm{x}) - (1 - t(\mathrm{x}))\mathbf{A}(\mathrm{x})}{t(\mathrm{x})}, \quad (3)$$

given the input hazy image $\mathbf{I}(\mathrm{x})$, the transmission matrix $t(\mathrm{x})$ replicated for each channel, and the non-uniform atmospheric illumination map $\mathbf{A}(\mathrm{x})$.

Training procedure

We train the encoder-decoder network by employing the multi-scale training proposed in [9] in order to both consider the semantics of the whole input image and to impose coherence among local structures. The network is optimized by minimizing the following loss:

$$\mathscr{L} = \lambda_1 \mathscr{L}_{cycle} + \lambda_2 \mathscr{L}_{\mathbf{A}} + \lambda_3 \mathscr{L}_P, \quad (4)$$

where λ_1, λ_2 and λ_3 are three coefficients weighting the contribute of each loss. $\mathscr{L}_{cycle}$ is the cycle consistency loss [25] (see Fig. 2) which enforces the bond among the translations $J \to \hat{I}$ and $I \to \hat{J}$ (we remove x for the sake of simplicity). The loss is defined as $\mathscr{L}_{cycle} = |\mathbf{J} - \hat{\mathbf{J}}| + |\mathbf{I} - \hat{\mathbf{I}}|$, where the first term represents the mean absolute error (MAE) between the pixels of the ground-truth haze-free image $\mathbf{J}$ and the corresponding dehazed image $\hat{\mathbf{J}}$, while the second term measures the MAE between the pixels of the input hazy image $\mathbf{I}$ and the inverted hazy image, $\hat{\mathbf{I}}$. $\mathscr{L}_{\mathbf{A}}$ minimizes the MAE between the dark channel prior of [14] and the predicted atmospheric illumination map $\mathbf{A}$. Given that the dark channel prior provides the global image illumination represented as an RGB triplet, we replicate it along spatial dimensions in order to be compared with the predicted atmospheric illumination. Finally, the perceptual loss $\mathscr{L}_P$ [16] estimates the mean square error between the features extracted from both the ground-truth haze-free image $\mathbf{J}$ and the corresponding enhanced haze-free one $\hat{\mathbf{J}}$ using a VGG19 trained on ImageNet dataset.

Table 1: **Overview of the considered databases.** The columns report: the number of images, whether images are captured indoor "in" or outdoor "out", if image pairs are available (because the haze-free reference image is given), and finally if the haze is real or synthetically generated.

Dataset	#Img	Context	Paired	Haze
Dense-Haze [3]	125	in/out	✓	real
I-HAZE [2]	35	in	✓	real
IVC Waterloo [19]	25	in/out		real
O-HAZE [4]	45	out	✓	real
P-Haze [9]	17125	in/out	✓	synth
RESIDE-HSTSr [18]	10	out		real
RESIDE-HSTSs [18]	10	out	✓	synth

Experimental Setup

In this section, we describe the datasets along with the settings used for experiments. Then, quantitative and qualitative results achieved by the proposed method are compared with state-of-the-art dehazing algorithms.

Datasets

We consider several datasets to evaluate the effectiveness of the proposed method in enhancing a vast range of semantic concepts and scenes. More in detail we take into account both synthetic and real haze databases containing indoor and outdoor images. The databases we include in the experiments are the following: the two databases (I-HAZE and O-HAZE) used for the NTIRE2018-Dehazing Challenge [1]; the Dense-Haze database introduced for the NTIRE2019-Dehazing Challenge [5]; a huge synthetic database proposed in [9]; a subset of the RESIDE database [18].

Parameter settings

We implement the proposed method using the PyTorch package [21]. The proposed model is trained on a workstation equipped with an NVIDIA Titan X Pascal GPU on the P-Haze dataset. The degree D of the polynomial used for the expansion is equal to 3, and the loss weights values are the following $\lambda_1 = \lambda_2 = 0.5$ and $\lambda_3 = 1$. All the convolutional weights are initialized with the method proposed in [15], while all the biases are set equal to zero.

Table 2: Quantitative comparison with state-of-the-art methods on paired image databases in terms of PSNR and SSIM.

Method	Dense-Haze PSNR↑/SSIM↑	I-Haze PSNR↑/SSIM↑	O-Haze PSNR↑/SSIM↑	RESIDE-HSTSs PSNR↑/SSIM↑
AOD-Net (ICCV2017) [17]	10.84/0.4654	15.53/0.7622	15.46/0.6076	20.55/0.8973
DehazeNet (TIP2016) [11]	9.48/0.4595	14.58/0.6915	16.65/0.6397	**24.48/0.9153**
MSCNN (ECCV2016) [22]	9.82/0.4629	15.45/0.7287	17.53/0.6773	18.64/0.8168
HR-Dehazer (CVPRW2019) [9]	**16.19/0.6010**	13.60/0.7928	21.44/0.6678	17.36/0.8589
Ours	15.37/0.5637	**22.41/0.8457**	**21.58/0.7051**	18.19/0.8389

We train the model from scratch by using Adam optimizer with a fixed learning rate of 0.0001, a batch-size equal to 1, and a momentum term of 0.5 for a total of 30 epochs. Hyper-parameters for fine-tuning the model on other datasets are the same used in training apart from the learning rate which corresponds to 1e-5 and the number of epochs equal to 100.

Results

Table 2 reports the results on paired databases in terms of Peak Signal to Noise Ratio (PSNR) and the Structural Similarity (SSIM) [23] index, which are the commonly employed metrics to evaluate dehazing algorithms on paired images. The proposed dehazing method outperforms the other ones on both the high-resolution image databases introduced for the NTIRE-2018 challenge. Especially on the I-HAZE, we obtained a PSNR 7dB higher than the second one which is the AOD-Net. Our method seems to be effective even if the density of haze is very high, in fact it ranks second on the Dense-Haze dataset. The worse performance on RESIDE-HSTSs is motivated due to the low-resolution of images.

Figure 3 shows enhanced images for all the considered dehazing methods. For paired images, we want to highlight that images dehazed with the proposed method are sharper than the others. Visibility is increased for all the databases apart from the Dense-Haze, where some image regions are completely lost due to the intense haze of the input hazy-image. In this sense the HR-Dehazer recovers more details thanks to the fact that it is an image-to-image method and so it can generate contents. On unpaired images the proposed method looks very powerful for removing haze and also does not distort colors.

Conclusions

In this work we investigated the use of a physical model for the task of haze removal. We showed that by introducing cycle consistency it is possible to reduce the reconstruction error and obtain more meaningful transmission maps. A logical continuum of this work can be the exploitation of the cycle-consistency for the use of this system on unpaired image datasets. To tackle the lack of a ground-truth, also the investigation of a no-reference image quality metric [8] should be investigated for measuring quantitatively the performances and thus the reconstruction quality. Other target atmosphere light priors could also be investigated.

Acknowledgments

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan X Pascal GPU used for this research.

References

[1] Cosmin Ancuti, Codruta O Ancuti, and Radu Timofte. Ntire 2018 challenge on image dehazing: Methods and results. In *CVPR-W*, pages 891–901. IEEE, 2018.

[2] Cosmin Ancuti, Codruta O Ancuti, Radu Timofte, and Christophe De Vleeschouwer. I-haze: a dehazing benchmark with real hazy and haze-free indoor images. In *ACIVS*, pages 620–631. Springer, 2018.

[3] Codruta O. Ancuti, Cosmin Ancuti, Mateu Sbert, and Radu Timofte. Dense haze: A benchmark for image dehazing with dense-haze and haze-free images. In *ICIP*, pages 1014–1018. IEEE, 2019.

[4] Codruta O Ancuti, Cosmin Ancuti, Radu Timofte, and Christophe De Vleeschouwer. O-haze: a dehazing benchmark with real hazy and haze-free outdoor images. In *CVPR-W*, pages 754–762. IEEE, 2018.

[5] Codruta O Ancuti, Cosmin Ancuti, Radu Timofte, Luc Van Gool, Lei Zhang, and Ming-Hsuan Yang. Ntire 2019 challenge on image dehazing: Methods and results. In *CVPR-W*, pages 891–901. IEEE, 2019.

[6] Yuval Bahat and Michal Irani. Blind dehazing using internal patch recurrence. In *ICCP*, pages 1–9. IEEE, 2016.

[7] Dana Berman, Shai Avidan, et al. Non-local image dehazing. In *CVPR*, pages 1674–1682. IEEE, 2016.

[8] Simone Bianco, Luigi Celona, Paolo Napoletano, and Raimondo Schettini. On the use of deep learning for blind image quality assessment. *Signal, Image and Video Processing*, 12(2):355–362, 2018.

[9] Simone Bianco, Luigi Celona, Flavio Piccoli, and Raimondo Schettini. High-resolution single image dehazing using encoder-decoder architecture. In *CVPR-W*, pages 0–0, 2019.

[10] Simone Bianco, Claudio Cusano, Flavio Piccoli, and Raimondo Schettini. Artistic photo filter removal using convolutional neural networks. *Journal of Electronic Imaging*, 27(1):1 – 14, 2017.

[11] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11):5187–5198, 2016.

[12] Raanan Fattal. Dehazing using color-lines. *ACM Transactions on Graphics (TOG)*, 34(1):13, 2014.

[13] Adrian Galdran, Aitor Alvarez-Gila, Alessandro Bria, Javier Vazquez-Corral, and Marcelo Bertalmío. On the duality between retinex and image dehazing. In *CVPR*, pages 8212–8221. IEEE, 2018.

[14] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2341–2353, 2011.

[15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *ICCV*, pages 1026–1034, 2015.

[16] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *ECCV*, pages 694–711. Springer, 2016.

[17] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. Aod-net: All-in-one dehazing network. In *ICCV*, volume 1, page 7. IEEE, 2017.

[18] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1):492–505, 2019.

[19] Kede Ma, Wentao Liu, and Zhou Wang. Perceptual evaluation of single image dehazing algorithms. In *ICIP*, pages 3600–3604. IEEE, 2015.

[20] Earl J McCartney. Optics of the atmosphere: scattering by

Figure 3: **Qualitative comparison between state-of-the-art methods and the proposed one.** Results are divided in paired and unpaired image databases.

molecules and particles. *John Wiley and Sons, Inc., 421 p.*, 1976.

[21] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. 2017.

[22] Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang. Single image dehazing via multi-scale convolutional neural networks. In *ECCV*, pages 154–169. Springer, 2016.

[23] Zhou Wang, Alan C Bovik, Hamid R Sheikh, Eero P Simoncelli, et al. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[24] He Zhang, Vishwanath Sindagi, and Vishal M Patel. Multi-scale single image dehazing using perceptual pyramid deep network. In *CVPR-W*, pages 902–911, 2018.

[25] Tinghui Zhou, Philipp Krahenbuhl, Mathieu Aubry, Qixing Huang, and Alexei A Efros. Learning dense correspondence via 3d-guided cycle consistency. In *CVPR*, pages 117–126. IEEE, 2016.

Author Biography

Simone Bianco *is an Associate Professor of computer science with the Department of Informatics, Systems and Communication (DISCo), University of Milano - Bicocca. His current research interests include computer vision, machine learning, optimization algorithms, and color imaging.*

Luigi Celona *is a Post-Doctoral Fellow with the DISCo, University of Milano - Bicocca. His current research interests focus on image analysis and classification, machine learning, and face analysis.*

Flavio Piccoli *is a Post-doctoral Fellow with the DISCo, University of Milano-Bicocca. He is focusing on image enhancement, image generation and anomaly detection for automatic industrial quality inspection.*