SynFog: A Photo-realistic Synthetic Fog Dataset based on End-to-end Imaging Simulation for Advancing Real-World Defogging in Autonomous Driving

Supplementary Material

In this supplementary document, we provide additional details on our foggy image simulation method and more evaluations to complete the main paper.

A. Analysis of ASM

$$E(d,\lambda) = \underbrace{E_o(\lambda)e^{-\beta(\lambda)d}}_{transmission} + \underbrace{L_h(\infty,\lambda)\left(1-e^{-\beta(\lambda)d}\right)}_{airlight}. \tag{1}$$

The second term of Eq. (1) represents the airlight caused by the scattering of environmental illumination (direct sunlight, diffuse sky light and light reflected by the ground) by particles in the atmosphere. Fig. 1 describes the derivation of the airlight [11].

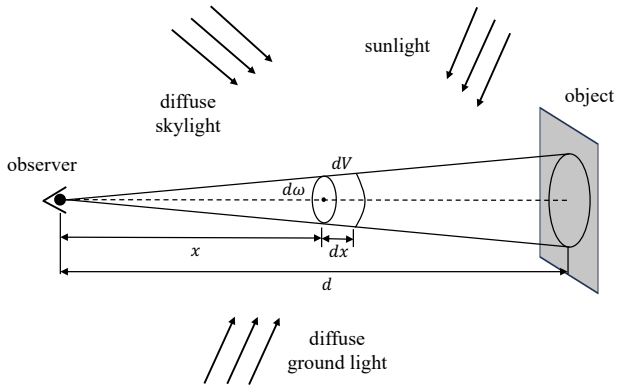

Figure 1. Airlight derivation. The cone of atmosphere between an observer and an object can be considered as a source of ambient illumination, termed airlight [11].

Specifically, if we define an infinitesimal volume of fog as $dV = d\omega x^2 dx$ and let k account for the exact nature of the illumination, the intensity of dV can be defined as:

$$dI(x,\lambda) = dVk\beta(\lambda) = d\omega x^2 dx k\beta(\lambda). \tag{2}$$

With the assumption that the linear dimensions of the object are smaller in comparison to its distance from the sensor [16], the infinitesimal volume can be treated as a point light source, and the irradiance it produces can be expressed as:

$$dE(x,\lambda) = \frac{dI(x,\lambda)e^{-\beta(\lambda)x}}{x^2}. \tag{3}$$

After obtaining radiance dL expressed in Eq. (4) from irradiance dE, we can get the final airlight L in Eq. (5) by integrating dL over the range of $x = 0$ to $x = d$.

$$dL(x,\lambda) = \frac{dE(x,\lambda)}{d\omega} = \frac{dI(x,\lambda)e^{-\beta(\lambda)x}}{d\omega x^2}. \tag{4}$$

$$L(d,\lambda) = k\left(1-e^{-\beta(\lambda)d}\right), \tag{5}$$

where $k = L_h(\infty,\lambda)$ represents the radiance of the sky at infinity. Based on the derivation of airlight, it is obvious that the construction of ASM [11] does not account for the effects of indirect illumination or active light sources on the scattering medium as shown in Eq. (2).

B. Construction of SynFog dataset

For the rendering of foggy scenes, we select 16 different high dynamic range sky maps for the development of SynFog as shown in Fig. 2. By incorporating these sky maps, we are able to create foggy scenes that closely resemble the real-world conditions under which fog typically occurs. Additional examples from SynFog are shown in Fig. 9.

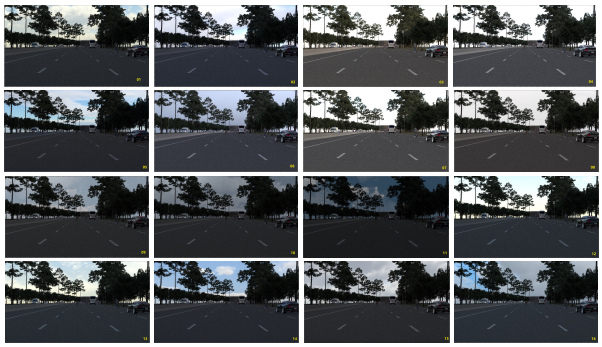

Figure 2. HDR sky maps used in our foggy scene rendering.

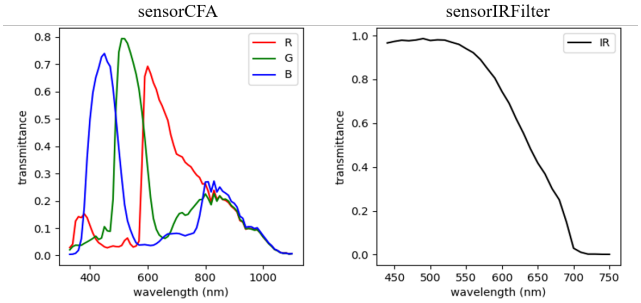

Figure 3. Characteristics of the sensor CFA (color filter array) and sensor IRFilter in our sensor simulation.

Table 1. Quantitative comparisons of SOTA defogging methods on SynFog dataset.

Method	SynFog (Ours)					RESIDE-indoor [8]	
	PSNR ↑	SSIM ↑	LPIPS [17] ↓	NIQE [10] ↓	FADE [2] ↓	PSNR ↑	SSIM ↑
(TPAMI'10) DCP [6]	18.30	0.540	0.367	4.202	0.669	16.62	0.818
(TIP'15) CAP [18]	17.35	0.364	0.418	4.289	**0.311**	18.96	0.815
(ICCV'19) GridNet [9]	26.98	0.851	**0.163**	5.620	0.980	32.16	0.984
(CVPR'20) MSBDN [4]	22.99	0.542	0.380	6.856	0.412	33.67	0.985
(AAAI'20) FFA-Net [12]	25.14	0.836	0.258	4.147	1.179	36.39	0.989
(CVPR'21) AECR-Net [15]	26.11	0.855	0.164	**3.973**	0.996	**37.17**	0.990
(TIP'23) DehazeFormer [14]	**27.96**	**0.923**	0.232	4.454	1.428	36.82	**0.992**

Table 2. Quantitative test results on corresponding synthetic datasets.

Datasets	Foggy Cityscapes [13]		Virtual KITTI [5]		SynFog (Ours)	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
AECR-Net [15]	33.9113	0.9873	29.4138	0.8761	26.1066	0.8551
DehazeFormer [14]	34.6294	0.9893	31.8525	0.9015	27.9552	0.9228

Figure 4. Comparison of SynFog with other synthetic datasets in the driving field. The second column displays the result after regional contrast stretching of the rectangular region, revealing that only SynFog exhibits consistent noise characteristics with those found in real captured foggy images.

Fig. 3 illustrates the actual sensor characteristics used in our sensor simulation, which includes the transmittance of the Color Filter Array (CFA) and the transmittance of the Infrared Filter. Fig. 4 shows the comparison of SynFog and the other two synthetic fog datasets after the enhancement of regional contrast stretching.

C. Fog Chamber Calibration

To ensure consistency between the virtual and real fog chamber, we calibrate the spectra of the light sources and the reflection spectra of the blackout cloth in the real fog chamber. The corresponding curves are depicted in Fig. 5. The color analysis under clear condition of the 24 patches on the color checker is shown in Fig. 6.

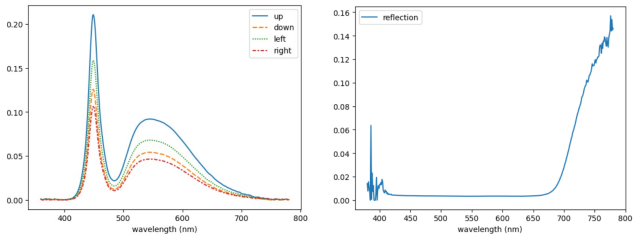

Figure 5. Calibration of the light sources and blackout cloth in the fog chamber.

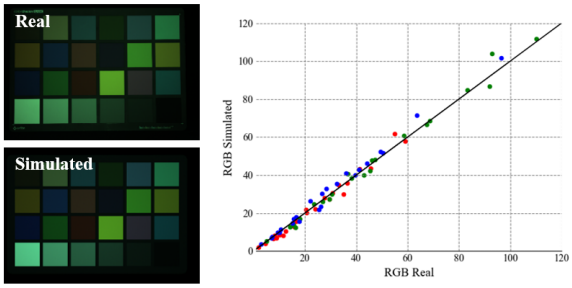

Figure 6. Color analysis under clear condition. In the scatter plot, the x-coordinates represent the RGB values from the real captured color patch, while the y-coordinates represent the RGB values from the simulated color patch. The colors red, green, and blue correspond to the three channels of the images, respectively.

D. Additional Experimental Results

Tab. 1 presents the test results for 7 mainstream defogging methods on both SynFog dataset and the RESIDE dataset [8]. The lower metrics achieved on SynFog compared to RESIDE indicate the challenges posed by SynFog and the potential for further improvement in existing defogging methods when applied to more realistic fog datasets.

Tab. 2 displays the test results of AECR-Net [15] and DehazeFormer [14] on their respective test sets of the training data. It can be seen that both models deliver strong performance on the corresponding test sets of Foggy Cityscapes [13] and Virtual KITTI [5], thus validating the effectiveness of the training process. The test results on real-world foggy images of these models are shown in Fig. 7. Only models trained on SynFog demonstrate superior generalization performance.

Table 3. Ablation Study using DehazeFormer.

Datasets	PSNR	SSIM
Foggy Cityscapes (Full)	34.6294	0.9893
Foggy Cityscapes (Part)	33.3741	0.9861
Virtual KITTI (Full)	31.8525	0.9015
Virtual KITTI (Part)	31.4001	0.8971
SynFog$-\beta$	28.5248	0.9428
SynFog (Ours)	27.9552	0.9228

Tab. 3 presents the test results of DehazeFormer [14] on the respective test sets of the training data when trained on the following conditions:

- Full Foggy Cityscapes and Part Foggy Cityscapes,
- Full Virtual KITTI and Part Virtual KITTI,
- SynFog-β and SynFog.

In the first two sets of experiments, we randomly select a training dataset of the same size as SynFog from the original one and use it as the Part dataset. The results show that the model trained on the Part dataset has poorer performance than the one trained on the Full dataset, on the test sets of both Foggy Cityscapes [13] and Virtual KITTI [5]. Moreover, the model trained on the Full dataset has worse generalization ability than that of SynFog, even when datasets are with larger scale. This highlights the advantages of our SynFog dataset compared with the other two datasets.

References

[1] Mario Bijelic, Tobias Gruber, Fahim Mannan, Florian Kraus, Werner Ritter, Klaus Dietmayer, and Felix Heide. Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 11682–11692, 2020. 4

[2] Lark Kwon Choi, Jaehee You, and Alan Conrad Bovik. Referenceless prediction of perceptual fog density and perceptual image defogging. *IEEE Transactions on Image Processing*, 24(11):3888–3901, 2015. 2

[3] Dengxin Dai, Christos Sakaridis, Simon Hecker, and Luc Van Gool. Curriculum model adaptation with synthetic and real data for semantic foggy scene understanding. *International Journal of Computer Vision*, 128:1182–1204, 2020. 4

[4] Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, Xinyi Zhang, Fei Wang, and Ming-Hsuan Yang. Multi-scale boosted dehazing network with dense feature fusion. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2157–2167, 2020. 2

[5] Adrien Gaidon, Qiao Wang, Yohann Cabon, and Eleonora Vig. Virtual worlds as proxy for multi-object tracking analysis. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 4340–4349, 2016. 2, 3

[6] 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, 2010. 2

[7] Glenn Jocher, Ayush Chaurasia, Alex Stoken, Jirka Borovec, Yonghye Kwon, Jiacong Fang, Kalen Michael, Diego Montes, Jebastin Nadar, Piotr Skalski, et al. ultralytics/yolov5: v6. 1-tensorrt, tensorflow edge tpu and openvino export and inference. *Zenodo*, 2022. 4

[8] 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, 2018. 2, 3

[9] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen. Griddehazenet: Attention-based multi-scale network for image dehazing. In *IEEE International Conference on Computer Vision*, pages 7314–7323, 2019. 2

[10] Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. Making a "completely blind" image quality analyzer. *IEEE Signal Processing Letters*, 20(3):209–212, 2012. 2

[11] Srinivasa G Narasimhan and Shree K Nayar. Vision and the atmosphere. *International Journal of Computer Vision*, 48: 233–254, 2002. 1

[12] Xu Qin, Zhilin Wang, Yuanchao Bai, Xiaodong Xie, and Huizhu Jia. Ffa-net: Feature fusion attention network for single image dehazing. In *AAAI*, pages 11908–11915, 2020. 2

[13] Christos Sakaridis, Dengxin Dai, and Luc Van Gool. Semantic foggy scene understanding with synthetic data. *International Journal of Computer Vision*, 126:973–992, 2018. 2, 3

[14] Yuda Song, Zhuqing He, Hui Qian, and Xin Du. Vision transformers for single image dehazing. *IEEE Transactions on Image Processing*, 32:1927–1941, 2023. 2, 3, 4

[15] Haiyan Wu, Yanyun Qu, Shaohui Lin, Jian Zhou, Ruizhi Qiao, Zhizhong Zhang, Yuan Xie, and Lizhuang Ma. Contrastive learning for compact single image dehazing. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 10551–10560, 2021. 2, 3, 4

[16] Lin Zhang, Anqi Zhu, Shiyu Zhao, and Yicong Zhou. Simulation of atmospheric visibility impairment. *IEEE Transactions on Image Processing*, 30:8713–8726, 2021. 1

Figure 7. Real-world test results of AECR-Net [15]. (a)-(c) are from Foggy Zurich dataset [3], while (d) is from Seeing Through Fog dataset [1]. The defogged images are further processed by a pretrained YOLOv5 [7] detection model. Only the model trained on SynFog demonstrates superior generalization performance on real-world foggy images with fewer artifacts and improved detection accuracy.

Figure 8. Real-world test results of DehazeFormer [14]. (a) is from Seeing Through Fog dataset [1], while (b) is from Foggy Zurich dataset [3]. We also compare the generalization ability of the model trained on SynFog-β. The results are consistent with the findings in Fig. 7.

Figure 9. Additional example images from SynFog dataset with various fog density levels.

[17] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 586–595, 2018. 2

[18] Qingsong Zhu, Jiaming Mai, and Ling Shao. A fast single image haze removal algorithm using color attenuation prior. *IEEE Transactions on Image Processing*, 24(11):3522–3533, 2015. 2