NEURAL INFORMATION PROCESSING SYSTEMS

Learning to dehaze with polarization

Chu Zhou[1], Minggui Teng[1], Yufei Han[3], Chao Xu[1], Boxin Shi[1,2*]

[1]Peking University [2]Beijing Academy of Artificial Intelligence [3]Beijing University of Posts and Telecommunications

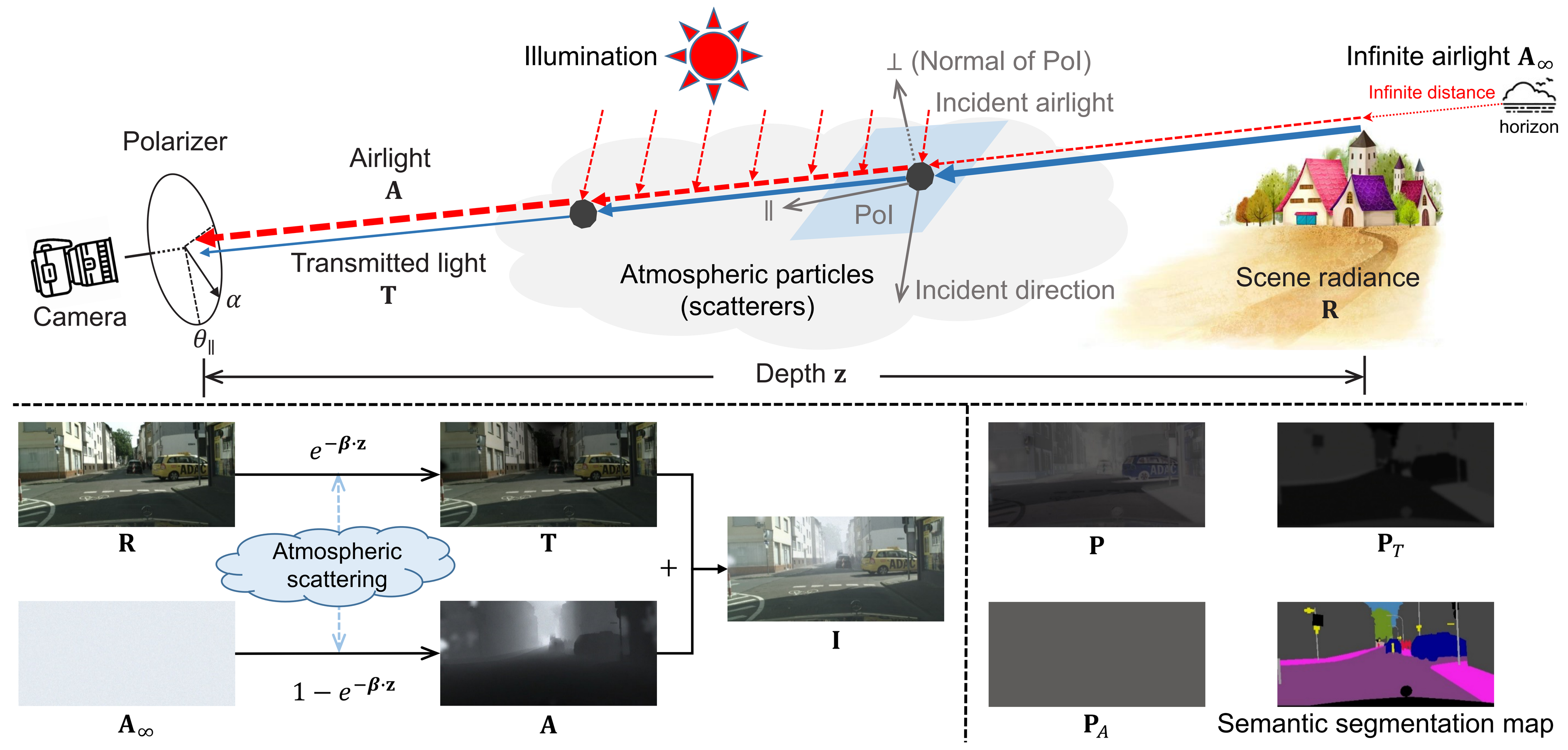
zhou_chu@pku.edu.cn, shiboxin@pku.edu.cn *Corresponding author

北京大學 PEKING UNIVERSITY

Contributions

- A generalized physical formation model of hazy images
  - taking into account the polarization effects of both transmitted light and airlight, along with the spatially-variant real-world scattering.
- A robust polarization-based dehazing pipeline
  - without the requirement of specific clues, by adopting deep learning to estimate necessary physical parameters.
- A two-stage neural network
  - making full use of semantic and contextual information to handle the spatially-variant real-world scattering to improve the clarity of original scene radiance recovery.

Physical image formation model



- Directly capturing: $\mathbf{I} = \mathbf{T} + \mathbf{A} = \mathbf{R}\cdot e^{-\beta\cdot \mathbf{z}} + \mathbf{A}_\infty \cdot (1 - e^{-\beta\cdot\mathbf{z}})$.
- Degree of polarization (DoP): $\mathbf{P} \triangleq \frac{\mathbf{I}^\perp - \mathbf{I}^\parallel}{\mathbf{I}}$ (spatially-variant), $\mathbf{P}_T \triangleq \frac{\mathbf{T}^\perp - \mathbf{T}^\parallel}{\mathbf{T}}$ (approximately uniform in the same semantic segment), $\mathbf{P}_A \triangleq \frac{\mathbf{A}^\perp - \mathbf{A}^\parallel}{\mathbf{A}}$ (spatially-uniform), where $\mathbf{I} = \mathbf{I}^\perp + \mathbf{I}^\parallel, \mathbf{T} = \mathbf{T}^\perp + \mathbf{T}^\parallel, \mathbf{A} = \mathbf{A}^\perp + \mathbf{A}^\parallel$.
- Placing a polarizer with polarization angle α: $\mathbf{I}_\alpha = \mathbf{T}_\alpha + \mathbf{A}_\alpha$, where $\mathbf{I}_\alpha = \frac{\mathbf{I}\cdot(1-\mathbf{P}\cdot\cos(2(\alpha-\theta_\parallel)))}{2}, \mathbf{T}_\alpha = \frac{\mathbf{T}\cdot(1-\mathbf{P}_T\cdot\cos(2(\alpha-\theta_\parallel)))}{2}, \mathbf{A}_\alpha = \frac{\mathbf{A}\cdot(1-\mathbf{P}_A\cdot\cos(2(\alpha-\theta_\parallel)))}{2}$.
- The relationship among **I**, **T**, and **A**: $\mathbf{I}\cdot\mathbf{P} = \mathbf{T}\cdot\mathbf{P}_T + \mathbf{A}\cdot\mathbf{P}_A$.

* β is the scattering coefficient, the subscript $\parallel$ ($\perp$) means the component is parallel (perpendicular) to the PoI, and $\theta_\parallel$ denotes the orientation of the polarizer for best transmission of the component parallel to the PoI.

Pipeline & network

Goal

- To restore the scene radiance **R** using three polarized images $\mathbf{I}_{\alpha^{(i)}}$ $(i = 1,2,3)$ captured at the same view with different polarization angles $\alpha^{(i)}$ $(i = 1,2,3)$.

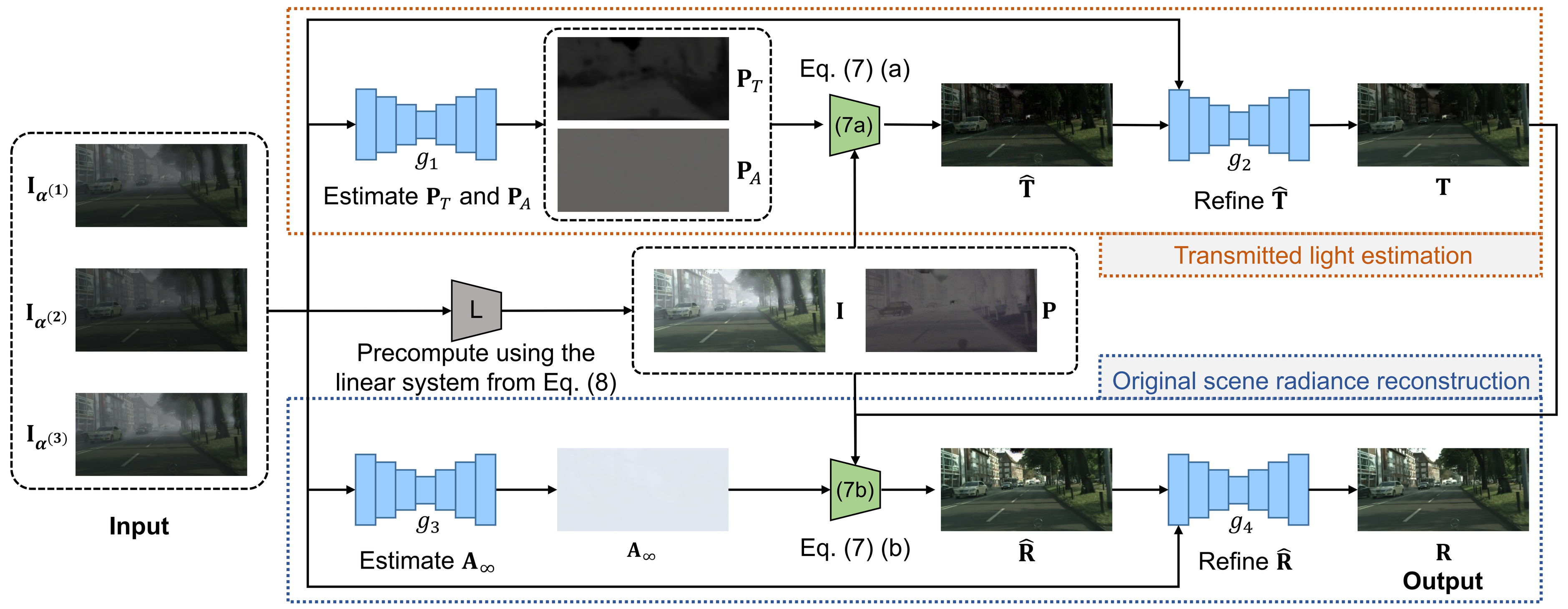
Polarization-based dehazing pipeline

A two-stage dehazing pipeline based on the physical image formation model:

(Eq. (7) of the paper) (a) $\mathbf{T} = \frac{\mathbf{P}\cdot\mathbf{I} - \mathbf{I}\cdot\mathbf{P}_A}{\mathbf{P}_T - \mathbf{P}_A}$ (b) $\mathbf{R} = \frac{\mathbf{T}\cdot\mathbf{A}_\infty}{\mathbf{A}_\infty - (\mathbf{I}-\mathbf{T})}$,

which means that

(a) once **I**, **P**, $\mathbf{P}_T$, and $\mathbf{P}_A$ are available, calculating **T** becomes well-posed;
(b) once **I**, **T**, and $\mathbf{A}_\infty$ are available, calculating **R** also becomes well-posed.



* **I** and **P** can be directly calculated by $\mathbf{I}_{\alpha^{(i)}}$ $(i = 1,2,3)$ using a linear system derived from the following equation (Eq. (8) of the paper):

$$\mathbf{I}_\alpha = \left\langle \left[\frac{1}{2} \quad \frac{-\cos(2\alpha)}{2} \quad \frac{-\sin(2\alpha)}{2}\right], [\mathbf{D}_1 \quad \mathbf{D}_2 \quad \mathbf{D}_3] \right\rangle,$$

where $\mathbf{D}_1 = \mathbf{I}$, $\mathbf{D}_2 = \mathbf{I}\cdot\mathbf{P}\cdot\cos(2\theta_\parallel)$ and $\mathbf{D}_3 = \mathbf{I}\cdot\mathbf{P}\cdot\sin(2\theta_\parallel)$.

Network Architecture

A two-stage neural network tailored to the polarization-based dehazing pipeline:

(a) **Transmitted light estimation** (two subnetworks g_1 and g_2):
- g_1: to estimate $\mathbf{P}_T$ and $\mathbf{P}_A$ for calculating $\hat{\mathbf{T}}$ (the coarse value of **T**);
- g_2: to refine $\hat{\mathbf{T}}$ for solving the numerical problem (happening when $\mathbf{P}_T \approx \mathbf{P}_A$) and handling the spatially-variant scattering.

(b) **Original scene radiance reconstruction** (two subnetworks g_3 and g_4):
- g_3: to estimate $\mathbf{A}_\infty$ for calculating $\hat{\mathbf{R}}$ (the coarse value of **R**);
- g_4: to refine $\hat{\mathbf{R}}$ for solving the numerical problem (happening when $\mathbf{T} \approx 0$) and handling the spatially-variant scattering.

Experiments

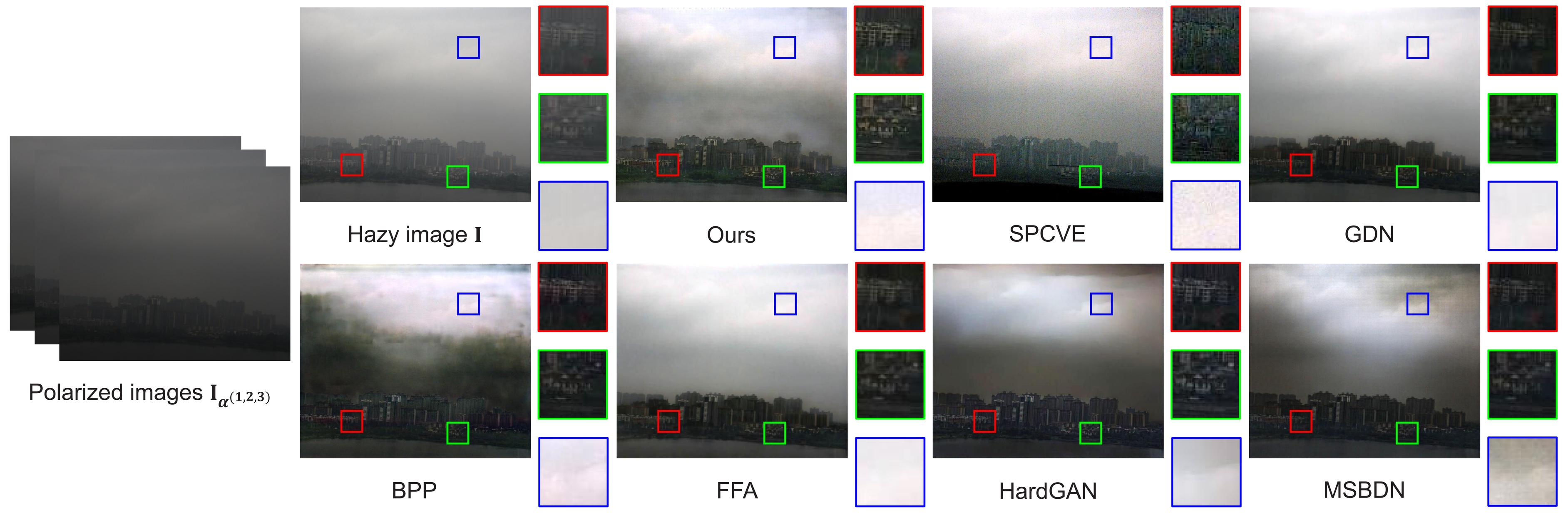
Quantitative results on synthetic data

	Ours	SPCVE	GDN	BPP	FFA	HardGAN	MSBDN
PSNR	**28.32**	15.94	26.54	24.93	26.84	26.22	26.94
MS-SSIM	**0.951**	0.521	0.928	0.915	0.934	0.928	0.932

Qualitative results on synthetic data

- P: PSNR M: MS-SSIM

Qualitative results on real data



- SPCVE: Namer *et al.*, Skyless polarimetric calibration and visibility enhancement. Optics Express, 2009.
- GDN: Liu *et al.*, GridDehazeNet: Attention-based multi-scale network for image dehazing. In Proc. of ICCV, 2019.
- BPP: Singh *et al.*, Single image dehazing for a variety of haze scenarios using back projected pyramid network. In Proc. of ECCVW, 2020.
- FFA: Qin *et al.*, FFA-Net: Feature fusion attention network for single image dehazing. In Proc. of AAAI, 2020.
- HardGAN: Deng *et al.*, HardGAN: A haze-aware representation distillation GAN for single image dehazing. In Proc. of ECCV, 2020.
- MSBDN: Dong *et al.*, Multi-scale boosted dehazing network with dense feature fusion. In Proc. of CVPR, 2020.