Polarimetric Neural Field via Unified Complex-Valued Wave Representation

Chu Zhou[1] Yixin Yang[2,3] Junda Liao[1,4] Heng Guo[5] Boxin Shi[2,3]* Imari Sato[1,4]*
[1]National Institute of Informatics, Japan
[2]State Key Laboratory of Multimedia Information Processing, School of Computer Science, Peking University, China
[3]National Engineering Research Center of Visual Technology, School of Computer Science, Peking University, China
[4]Graduate School of Information Science and Technology, University of Tokyo, Japan
[5]School of Artificial Intelligence, Beijing University of Posts and Telecommunications, China
zhou_chu@hotmail.com, {yangyixin93, shiboxin}@pku.edu.cn,
liao-junda@g.ecc.u-tokyo.ac.jp, guoheng@bupt.edu.cn, imarik@nii.ac.jp

Abstract

Polarization has found applications in various computer vision tasks by providing additional physical cues. However, due to the limitations of current imaging systems, polarimetric parameters are typically stored in discrete form, which is non-differentiable and limits their applicability in polarization-based vision. While current neural field methods have shown promise for continuous signal reconstruction, they struggle to model the intrinsic physical interdependencies among polarimetric parameters. In this work, we propose a physics-grounded representation scheme to represent polarimetric parameters as a unified complex-valued wave. Tailored to this scheme, we propose a tuning-free fitting strategy along with a lightweight complex-valued neural network, enabling property-preserved reconstruction. Experimental results show that our method achieves state-of-the-art performance and facilitates smooth polarized image rendering and flexible resolution adjustments.

1. Introduction

Polarization, alongside amplitude and phase, is a fundamental property of light and has shown promise in various computer vision tasks by providing additional physical cues. To capture the *polarimetric parameters* (including the total intensity (TI), degree of polarization (DoP), and angle of polarization (AoP)) of a scene, current polarimetric imaging methods typically require acquiring multiple polarized images, either by using linear polarizers or polarization cameras [43]. However, since the captured polarized images are stored in discrete form (*i.e.*, as pixels), the calculated polarimetric parameters are also discrete and non-differentiable, preventing smooth polarized image rendering and flexible resolution adjustments [29], thus limiting their applicability in tasks such as reflection removal [14, 22] and shape from polarization [6, 23]. In contrast, obtaining these parameters in continuous form would not only address these issues but also enable continuous optimization and improve memory efficiency. Therefore, reconstructing polarimetric parameters from discrete to continuous form is of practical significance.

Recent advances in neural fields [16, 21] have made progress in parameterizing varying physical quantities with neural networks. By leveraging various nonlinearities and architectural designs to fit directly, as shown in Fig. 1 (top left (a)), these methods enable reliable reconstruction of continuous signals from discrete data. However, when fitting multiple polarimetric parameters concurrently, they struggle with the wide range of spatial frequencies, introducing noticeable artifacts. To mitigate this issue, multi-band solutions [29] have been proposed. As shown in Fig. 1 (top left (b)), these methods decompose the polarimetric parameters into a series of bands and fit them sequentially. Nevertheless, existing methods, whether fitting in a direct manner or through a multi-band solution, still face several limitations:

- **Physics-unaware representation**: All methods treat different parameters identically, without explicitly considering their physical meanings and interdependencies.
- **Scene-dependent fitting**: Multi-band solutions require manually tuning the threshold of each band for each scene, hindering automation in the fitting process.
- **Parameter-heavy architecture**: Multi-band solutions naturally necessitate larger network capacity, leading to increased parameter count and model size.

We observe that different polarimetric parameters are inherently correlated through the Stokes parameters [12], meaning that a change in one inevitably affects the others. Interestingly, this interdependence resembles quantum entanglement [4], where components are intrinsically linked

---

*Corresponding authors.

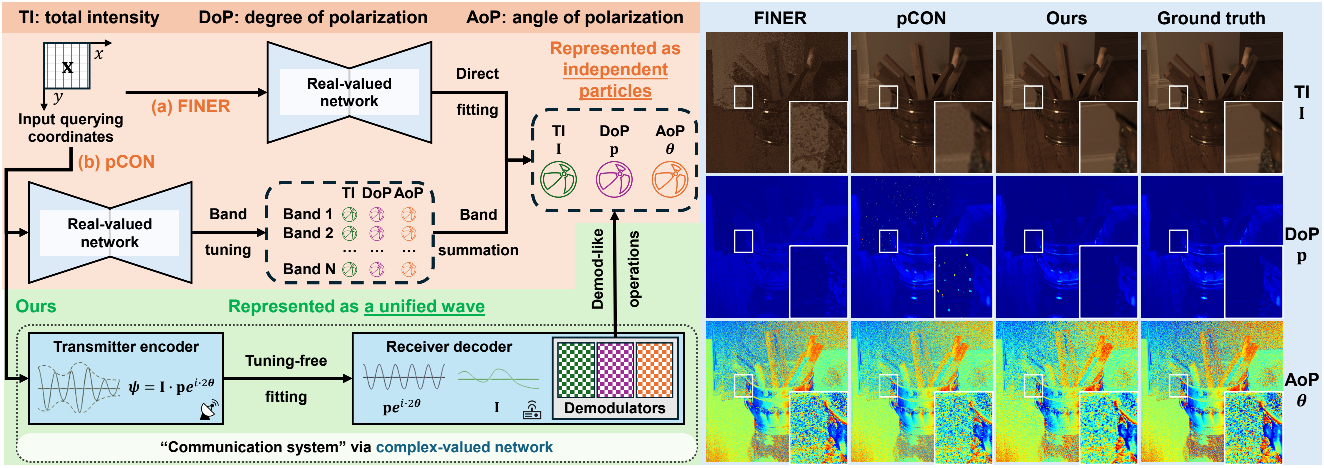

Figure 1. Top left: Current neural field methods represent the polarimetric parameters (including the total intensity (TI), degree of polarization (DoP), and angle of polarization (AoP)) as independent particles and use real-valued networks to fit them concurrently, either in a direct manner (*e.g.*, (a) FINER [21]) or through a multi-band solution (*e.g.*, (b) pCON [29]). Bottom left: In contrast, we represent the polarimetric parameters as a unified wave and utilize a complex-valued network to fit them in a tuning-free manner, modeling the fitting process as a "communication system" with a transmitter encoder and a receiver decoder. Right: The reconstruction results of FINER [21], pCON [29], and ours. Following previous works [19, 42], we visualize the DoP and AoP using color maps. Please zoom in for better details.

rather than independent. Inspired by this analogy, we propose to represent the polarimetric parameters as *a unified complex-valued wave* (Fig. 1 (bottom left)), instead of as *independent particles* (Fig. 1 (top left)). This representation scheme not only effectively captures the interdependencies among polarimetric parameters, but also aligns with the physical image formation model of polarized images. Tailored to this scheme, we introduce a tuning-free fitting strategy that fully decouples the fitting process from scene dependency. By modeling the fitting process as a "communication system" including a transmitter encoder and a receiver decoder, the wave is treated as a modulated signal, with demodulation-like operations employed to robustly retrieve each polarimetric parameter, similar to separating the carrier and message signals. Specifically, we design a lightweight complex-valued neural network to implement the fitting strategy efficiently, seamlessly adapting to the formulation of wave while achieving superior performance. To summarize, this paper contributes by demonstrating:

- **A physics-grounded representation scheme**, integrating the physical meanings and interdependencies of the polarimetric parameters into a unified wave.
- **A tuning-free fitting strategy**, fully decoupling the fitting process from scene dependency and enabling robust retrieval of each polarimetric parameter.
- **A lightweight complex-valued network**, seamlessly adapting to the formulation of wave with a compact architecture and superior performance.

Experimental results show that our method achieves state-of-the-art performance and facilitates smooth polarized image rendering and flexible resolution adjustments.

2. Related Work

Polarization-based vision. Polarization-based vision algorithms aim to conquer the performance bottleneck of image-based ones, by fully exploiting the additional physical cues encoded in polarimetric parameters. Some methods are designed for solving high-level vision problems in applications such as robotics (*e.g.*, transparent object segmentation for bin picking [9, 26]) and autonomous driving (*e.g.*, road scene understanding [18, 20]). Some methods focus on 3D vision, such as shape estimation [6, 23] and depth sensing [8, 37]. There are also methods focus on image enhancement, such as reflection removal [14, 22], image dehazing [33, 40], HDR reconstruction [41], and shadow removal [45].

Polarimetric imaging. Polarimetric imaging aims to obtain the polarimetric parameters of the scene by capturing polarized images. To improve the accuracy, various methods have been proposed to enhance the quality of captured polarized images through post-processing. They focus on explicitly mitigating degradations caused by demosaicing artifacts [30, 46], low-light noise [19, 42], motion blur [44], *etc*. Moving beyond post-processing, computational photography-based polarimetric imaging frameworks [43] have also been explored. However, these methods still produce discrete outputs, limiting their usage in continuous optimization.

Neural fields. A neural field is a field (*i.e.*, a varying physical quantity of coordinates) parameterized by a neural network, typically based on multilayer perceptrons (MLPs) [39]. Within MLPs, ReLU is the most commonly used nonlinearity, often combined with positional encoding [36]. This approach has been widely applied in novel view synthesis

[25, 27], shape representation [5, 28], and medical imaging [35, 38], *etc*. Recently, alternative nonlinearities have been introduced to improve signal encoding quality, including sinusoidal functions [34], Gaussian functions [31], complex Gabor wavelets [32], and variable-periodic functions (FINER) [21]. Additionally, novel architectural designs have been proposed to improve fitting efficiency, such as multi-scale block-coordinate decomposition [24], generalized superpixels (S-INR) [16], and spatially collaged Fourier bases [17]. While these methods tend to introduce noticeable artifacts when applied to polarimetric parameters, pCON [29], the most related work, proposes a multi-band solution to mitigate this issue. However, it requires scene-dependent manual band tuning and larger network capacity. It is worth noting that, while many neural field methods also incorporate polarization (*e.g*., using polarization for radiance decomposition [3], novel-view hyperspectral rendering [10], and geometry and material estimation [15]), their focus is on using polarization as guidance, rather than fitting the polarimetric parameters, which differs from our aim.

3. Method

3.1. Background and Problem Formulation

Polarimetric parameters: definition. Light propagates as a wave of electric and magnetic fields, with polarization describing the orientation of the electric field oscillations. While most natural light sources (*e.g*., sunlight) emit unpolarized light, various optical phenomena such as reflection, refraction, and scattering commonly induce partial polarization[1] in the light that reaches our eyes [12]. Denoting the total intensity (TI) of the light as $\mathbf{I}$, it can be decomposed into two components with different polarization properties:

$$\mathbf{I} = \mathbf{I}_u + \mathbf{I}_p, \tag{1}$$

where $\mathbf{I}_u$ is the unpolarized component (*i.e*., the electric field of $\mathbf{I}_u$ oscillates randomly and evenly) and $\mathbf{I}_p$ is the fully polarized component (*i.e*., the electric field $\mathbf{I}_p$ oscillates in a single plane with a constant orientation). When a linear polarizer with angle α is used to filter $\mathbf{I}$, the transmitted light intensity $\mathbf{I}_\alpha$ follows Malus' law [7]:

$$\mathbf{I}_\alpha = \frac{\mathbf{I}}{2} \cdot (1 - \mathbf{p} \cdot \cos(2(\alpha - \boldsymbol{\theta}))), \tag{2}$$

where $\mathbf{p} \in [0, 1]$ is the degree of polarization (DoP) and $\boldsymbol{\theta} \in [0, \pi]$ is the angle of polarization (AoP)[2]. DoP measures how much the light is polarized, which is defined as

$$\mathbf{p} = \frac{\mathbf{I}_p}{(\mathbf{I}_u + \mathbf{I}_p)} = \frac{\mathbf{I}_p}{\mathbf{I}}, \tag{3}$$

and AoP describes the orientation of the electric field of $\mathbf{I}_p$. Since $\mathbf{I}_\alpha$ can be rendered for any arbitrary α given $\mathbf{I}$, $\mathbf{p}$, and $\boldsymbol{\theta}$, these three physical quantities (TI, DoP, and AoP) are referred to as polarimetric parameters [29].

Polarimetric parameters: discrete acquisition. Since $\mathbf{I}_u$ and $\mathbf{I}_p$ are unknown, polarimetric parameters cannot be obtained directly. Instead, current polarimetric imaging methods estimate them by capturing multiple polarized images. Denoting $\mathbf{I}_{\alpha_{1,2,3,4}} \in \mathbb{R}^{3\times h\times w}$ as the polarized images captured at $\alpha_{1,2,3,4} = 0°, 45°, 90°, 135°$, the corresponding Stokes parameters $\mathbf{S}_{0,1,2} \in \mathbb{R}^{3\times h\times w}$ [12] are defined as

$$\begin{cases} \mathbf{S}_0 = \mathbf{I}_{\alpha_1} + \mathbf{I}_{\alpha_3} = \mathbf{I}_{\alpha_2} + \mathbf{I}_{\alpha_4} \\ \mathbf{S}_1 = \mathbf{I}_{\alpha_3} - \mathbf{I}_{\alpha_1}, \ \text{ and } \ \mathbf{S}_2 = \mathbf{I}_{\alpha_4} - \mathbf{I}_{\alpha_2} \end{cases}, \tag{4}$$

and the polarimetric parameters can be calculated using

$$\mathbf{I} = \mathbf{S}_0, \mathbf{p} = \frac{\sqrt{\mathbf{S}_1^2 + \mathbf{S}_2^2}}{\mathbf{S}_0} \ \text{ and } \ \boldsymbol{\theta} = \frac{1}{2}\arctan(\frac{\mathbf{S}_2}{\mathbf{S}_1}). \tag{5}$$

Thus, the polarimetric parameters are acquired indirectly and stored in discrete form similar to $\mathbf{I}_{\alpha_{1,2,3,4}}$ (*i.e*., $\mathbf{I}, \mathbf{p}, \boldsymbol{\theta} \in \mathbb{R}^{3\times h\times w}$), limiting their usage in continuous optimization.

Polarimetric parameters: continuous reconstruction. Our goal is to design a polarimetric neural field to reconstruct continuous polarimetric parameters from their discrete observations while preserving the polarization properties of light. Specifically, this involves concurrently fitting both TI, DoP, and AoP using a neural network $f(\mathbf{x}; \eta)$, where $\mathbf{x}$ represents the input querying coordinates, η denotes the network parameters, and the network output is expected to closely approximate the sampled values of TI, DoP, and AoP at $\mathbf{x}$. Due to the differentiable nature of the neural network, polarimetric parameters can be queried at any resolution, enabling inherently continuous storage.

3.2. Physics-Grounded Representation Scheme

Before fitting, it is essential to design a suitable representation scheme that helps the network to interpret the characteristics of the polarimetric parameters. As shown in Fig. 1 (top left), existing methods [16, 21, 29] adopt a straightforward way to represent $\mathbf{I}$, $\mathbf{p}$, and $\boldsymbol{\theta}$ as independent particles, treating them identically without explicitly considering their physical meanings and interdependencies. However, as indicated by Eq. (4) and Eq. (5), both $\mathbf{I}$ and $\mathbf{p}$ are related to $\mathbf{S}_0$, while $\mathbf{p}$ and $\boldsymbol{\theta}$ are related to $\mathbf{S}_{1,2}$. This reveals that $\mathbf{I}$, $\mathbf{p}$, and $\boldsymbol{\theta}$ are entangled with each other. As a proof of concept, we select two different scenes and calculate the mutual information (MI) [1] between various physical quantities, as illustrated

---

[1] In this paper, like pCON [29], we focus on linear polarization, as most natural optical phenomena do not induce circular polarization. For simplicity, we refer to linear polarization simply as polarization.

[2] In some works, Eq. (2) appears as $\mathbf{I}_\alpha = \mathbf{I} \cdot (1 + \mathbf{p} \cdot \cos(2(\boldsymbol{\theta} - \alpha)))/2$ [9, 29], which is equally valid but follows a different convention. We adopt the formulation used in previous works such as [33, 42]. Discussions about the conventions can be found in the supplementary material.

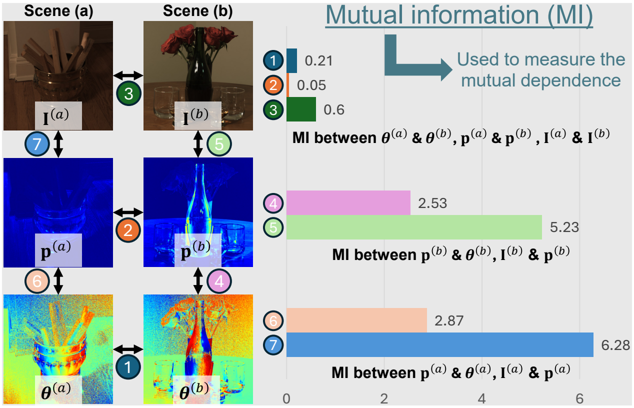

Figure 2. The mutual information between two different polarimetric parameters within the same scene is much higher than that between the same polarimetric parameter across different scenes, confirming the interdependencies among polarimetric parameters.

in Fig. 2. We can see that the MI between two different polarimetric parameters within the same scene is much higher than that between the same polarimetric parameter across different scenes, confirming the interdependencies among polarimetric parameters.

We observe that the interdependencies among polarimetric parameters bear similarities to quantum entanglement: in quantum mechanics, entangled particles should be represented as an inseparable whole (*i.e.*, a wave) instead of being described independently [4]. Inspired by this concept, as shown in Fig. 1 (bottom left), we propose to represent $\mathbf{I}$, $\mathbf{p}$, and $\boldsymbol{\theta}$ as a unified wave $\boldsymbol{\psi}$:

$$\boldsymbol{\psi}(\mathbf{I}, \mathbf{p}, \boldsymbol{\theta}) = \boldsymbol{z} e^{i\boldsymbol{\gamma}} = \mathbf{I} \cdot \mathbf{p} e^{i \cdot 2\boldsymbol{\theta}}, \tag{6}$$

where $\boldsymbol{z} = \mathbf{I} \cdot \mathbf{p}$ and $\boldsymbol{\gamma} = 2\boldsymbol{\theta}$ denote its amplitude and phase, and i is the imaginary unit satisfying $i^2 = -1$. This representation scheme offers several advantages. First, as shown in Eq. (6), the amplitude $\boldsymbol{z}$ directly corresponds to the fully polarized component of light ($\mathbf{I}_p$, see Eq. (1) and Eq. (3) for details), establishing a clear physical connection between TI and DoP. This explicit linkage enables the fitting process to effectively exploit their mutual information for improved accuracy. Additionally, the phase $\boldsymbol{\gamma}$ naturally encodes AoP as an angular quantity, inherently capturing its periodic nature, which aligns well with its physical meaning.

Here, we also show that the proposed representation scheme is consistent to the physical image formation model of polarized images. Reformulating $\mathbf{S}_{1,2}$ in Eq. (4) as the functions of $\mathbf{I}$, $\mathbf{p}$, and $\boldsymbol{\theta}$, we can obtain

$$\mathbf{S}_1 = \mathbf{I} \cdot \mathbf{p} \cdot \cos(2\boldsymbol{\theta}) \quad \text{and} \quad \mathbf{S}_2 = \mathbf{I} \cdot \mathbf{p} \cdot \sin(2\boldsymbol{\theta}). \tag{7}$$

Expanding Eq. (6) with Euler's formula and substituting Eq. (7) into it, we could get the following equation:

$$\boldsymbol{\psi}(\mathbf{I}, \mathbf{p}, \boldsymbol{\theta}) = \mathbf{I} \cdot \mathbf{p} \cdot (\cos(2\boldsymbol{\theta}) + i \cdot \sin(2\boldsymbol{\theta})) = \mathbf{S}_1 + i \cdot \mathbf{S}_2, \tag{8}$$

which demonstrates that $\boldsymbol{\psi}$ inherently follows the additive properties of the Stokes parameters. Therefore, in addition to sharing a similar form with a wave, $\boldsymbol{\psi}$ also adheres to the following rules:

Rule 1. *A linear combination of multiple beams of light can be represented as a linear combination of multiple waves shown in Eq. (6).*

Explanation. Using the superscripts j and $comb$ to denote the individual terms within the linear combination and the resulting combination, we can deduce

$$\begin{aligned} \sum_{j=1}^{n} w^{(j)} \boldsymbol{\psi}^{(j)} &= \sum_{j=1}^{n} w^{(j)} \mathbf{S}_1^{(j)} + i \cdot \sum_{j=1}^{n} w^{(j)} \mathbf{S}_2^{(j)} \\ &= \mathbf{S}_1^{(comb)} + i \cdot \mathbf{S}_2^{(comb)} = \boldsymbol{\psi}^{(comb)}, \end{aligned} \tag{9}$$

where $w^{(j)}$ is the weight of the j-th term. Here, the DoP and AoP of the resulting combination can be easily obtained by

$$\boldsymbol{\theta}^{(comb)} = \frac{\angle \boldsymbol{\psi}^{(comb)}}{2} \quad \text{and} \quad \mathbf{p}^{(comb)} = \frac{\|\boldsymbol{\psi}^{(comb)}\|}{\mathbf{I}^{(comb)}}, \tag{10}$$

where $\angle\cdot$ and $\|\cdot\|$ denote extracting the phase and magnitude respectively, and $\mathbf{I}^{(comb)} = \sum_{j=1}^{n} w^{(j)} \mathbf{I}^{(j)}$, which is consistent to the superposition of linear polarized light in the physical world.

Rule 2. *Using a polarizer to filter a beam of light corresponds to projecting its wave representation onto the state defined by the polarizer angle, and the outcome of this "observation" is the component aligned with the polarizer.*

Explanation. Expanding Eq. (2), we can get $\mathbf{I}_\alpha = \frac{1}{2}(\mathbf{I} - \mathbf{I} \cdot \mathbf{p} \cdot \cos(2(\boldsymbol{\theta} - \alpha)))$. Then, according to the fact that the cosine function can be expressed as the real part of a complex exponential, we can deduce

$$\mathbf{I}_\alpha = \frac{1}{2}(\mathbf{I} - \mathbf{I} \cdot \mathbf{p} \cdot \Re(e^{i \cdot 2(\boldsymbol{\theta} - \alpha)})) = \frac{1}{2}(\mathbf{I} - \Re(\boldsymbol{\psi} e^{-i \cdot 2\alpha})) \tag{11}$$

where $\Re(\cdot)$ denotes the operator for extracting the real part. From the above equation, we can see that the term $\boldsymbol{\psi} e^{-i \cdot 2\alpha}$ has a product form, corresponding to the "observation" produced by projecting $\boldsymbol{\psi}$ onto the direction defined by the polarizer angle α, which is consistent with the physics underlying real-world polarimetric imaging.

3.3. Tuning-Free Fitting Strategy

Due to the large domain gap among $\mathbf{I}$, $\mathbf{p}$, and $\boldsymbol{\theta}$, which introduces a wide range of spatial frequencies, methods based on direct fitting [16, 21] often result in noticeable artifacts, making the design of an effective fitting strategy challenging. To mitigate the artifacts, multi-band solutions [29] decompose the polarimetric parameters into multiple frequency bands and fit them sequentially. However, these methods treat band thresholds as hyperparameters that require manual tuning for each scene, which hinders automation in the fitting process.

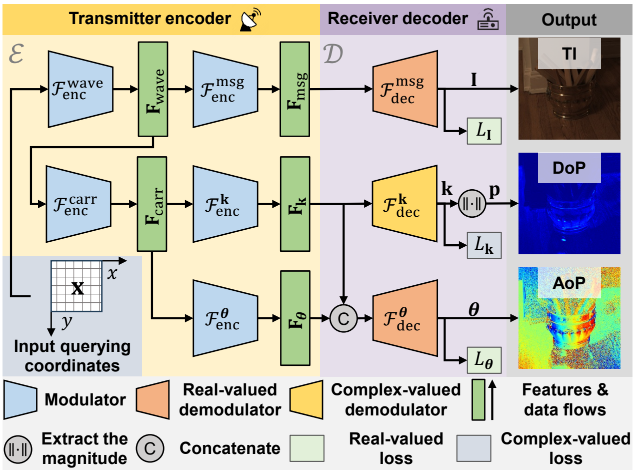

Figure 3. The proposed tuning-free fitting strategy tailored to our physics-grounded representation. By modeling the fitting process as a "communication system" including a transmitter encoder and a receiver decoder, it employs demodulation-like operations to separate the carrier and message signals for retrieving each polarimetric parameter, making full use of their interdependencies.

To eliminate scene dependency in the fitting process, we introduce a tuning-free strategy tailored to our physics-grounded representation, combining the simplicity of direct fitting with effective artifact suppression. Notably, we observe that Eq. (8) can be written as

$$\psi(\mathbf{I},\mathbf{p},\boldsymbol{\theta}) = \mathbf{I}\cdot(\mathbf{p}\cdot\cos(2\boldsymbol{\theta}) + i\cdot\mathbf{p}\cdot\sin(2\boldsymbol{\theta})) = \mathbf{I}\cdot\mathbf{k}, \quad (12)$$

where $\mathbf{k}$ is a complex quantity whose amplitude is confined within the unit circle, as $\mathbf{p} \in [0, 1]$. This implies that the structure of ψ resembles a modulated complex signal, with $\mathbf{I}$ acting as the message signal modulating the carrier signal $\mathbf{k} = \mathbf{p}\cdot\cos(2\boldsymbol{\theta}) + i\cdot\mathbf{p}\cdot\sin(2\boldsymbol{\theta}) = \mathbf{p}e^{i\cdot 2\boldsymbol{\theta}}$. Based on this observation, we propose to model the fitting process as a "communication system" including a transmitter encoder and a receiver decoder, and employ demodulation-like operations to separate the carrier and message signals for retrieving each polarimetric parameter, as shown in Fig. 3.

To fully exploit the interdependencies among the polarimetric parameters, we design the transmitter encoder as a top-down hierarchical structure. First, we use a modulator, $\mathcal{F}^{\text{wave}}_{\text{enc}}$, to map the input querying coordinates $\mathbf{x}$ into $\mathbf{F}_{\text{wave}}$, which stands for the global features of the wave ψ, encapsulating the structural and contextual information of the scene shared by both TI, DoP, and AoP [43]. Subsequently, from $\mathbf{F}_{\text{wave}}$, two additional modulator branches, $\mathcal{F}^{\text{msg}}_{\text{enc}}$ and $\mathcal{F}^{\text{carr}}_{\text{enc}}$, are adopted to separately encode the distinct information of the message signal $\mathbf{I}$ and the carrier signal $\mathbf{k}$ into the corresponding features, $\mathbf{F}_{\text{msg}}$ and $\mathbf{F}_{\text{carr}}$. Here, since the spatial frequency of AoP is typically much higher than that of the DoP [29], directly applying demodulation-like operations to $\mathbf{F}_{\text{carr}}$ to optimize both the magnitude and phase of $\mathbf{k}$ may introduce frequency interference between the DoP and AoP. To mitigate this, from $\mathbf{F}_{\text{carr}}$, we add another two modulator branches, $\mathcal{F}^{\mathbf{k}}_{\text{enc}}$ and $\mathcal{F}^{\boldsymbol{\theta}}_{\text{enc}}$, to further encode the low-frequency textures of $\mathbf{k}$ and the high-frequency details exclusive to $\boldsymbol{\theta}$ into the features $\mathbf{F}_{\mathbf{k}}$ and $\mathbf{F}_{\boldsymbol{\theta}}$, respectively. The whole transmitter encoder $\mathcal{E}$ can be written as

$$\begin{aligned}&\mathbf{F}_{\text{msg}}, \mathbf{F}_{\mathbf{k}}, \mathbf{F}_{\boldsymbol{\theta}} = \mathcal{E}(\mathbf{x}), \quad \text{where} \;\; \mathbf{F}_{\text{wave}} = \mathcal{F}^{\text{wave}}_{\text{enc}}(\mathbf{x}),\\ &\mathbf{F}_{\text{msg}} = \mathcal{F}^{\text{msg}}_{\text{enc}}(\mathbf{F}_{\text{wave}}), \mathbf{F}_{\text{carr}} = \mathcal{F}^{\text{carr}}_{\text{enc}}(\mathbf{F}_{\text{wave}}),\\ &\mathbf{F}_{\mathbf{k}} = \mathcal{F}^{\mathbf{k}}_{\text{enc}}(\mathbf{F}_{\text{carr}}), \;\; \text{and} \;\; \mathbf{F}_{\boldsymbol{\theta}} = \mathcal{F}^{\boldsymbol{\theta}}_{\text{enc}}(\mathbf{F}_{\text{carr}}).\end{aligned} \quad (13)$$

For the receiver decoder, we design it as a parallel three-branches structure to decode the features $\mathbf{F}_{\text{msg}}$, $\mathbf{F}_{\mathbf{k}}$ and $\mathbf{F}_{\boldsymbol{\theta}}$ into their respective physical quantities. Specifically, in the first two branches, since the message signal $\mathbf{I}$ is real while the carrier signal $\mathbf{k}$ is complex, we adopt a real-valued demodulator $\mathcal{F}^{\text{msg}}_{\text{dec}}$ and a complex-valued demodulator $\mathcal{F}^{\mathbf{k}}_{\text{dec}}$ to decode $\mathbf{F}_{\text{msg}}$ and $\mathbf{F}_{\mathbf{k}}$ into $\mathbf{I}$ and $\mathbf{k}$, respectively. The DoP $\mathbf{p}$ is then computed as $\mathbf{p} = \|\mathbf{k}\|$. In the third branch, rather than directly decoding $\mathbf{F}_{\boldsymbol{\theta}}$ into $\boldsymbol{\theta}$, we choose to make full use of both $\mathbf{F}_{\mathbf{k}}$ and $\mathbf{F}_{\boldsymbol{\theta}}$ by applying another real-valued demodulator $\mathcal{F}^{\boldsymbol{\theta}}_{\text{dec}}$ to their concatenation. This is because $\mathbf{k}$ inherently carries prior knowledge of $\boldsymbol{\theta}$ (*i.e.*, under ideal fitting, the phase of $\mathbf{k}$ should be $2\boldsymbol{\theta}$), suggesting that $\mathbf{F}_{\mathbf{k}}$ could contribute to enhancing the accuracy of $\boldsymbol{\theta}$. The whole receiver decoder $\mathcal{D}$ can be written as

$$\begin{aligned}&\mathbf{I} = \mathcal{F}^{\text{msg}}_{\text{dec}}(\mathbf{F}_{\text{msg}}), \mathbf{p} = \|\mathbf{k}\| = \|\mathcal{F}^{\mathbf{k}}_{\text{dec}}(\mathbf{F}_{\mathbf{k}})\|,\\ &\text{and} \;\; \boldsymbol{\theta} = \mathcal{F}^{\boldsymbol{\theta}}_{\text{dec}}(\texttt{concat}(\mathbf{F}_{\mathbf{k}}, \mathbf{F}_{\boldsymbol{\theta}})).\end{aligned} \quad (14)$$

With the above fitting strategy, we only need to focus on the module designs of the modulators and demodulators, getting rid of any scene-related hyperparameters. Besides, we can conveniently design loss functions for direct optimization, avoiding multi-band optimization.

3.4. Lightweight Complex-Valued Network

To unlock the full potential of our tuning-free fitting strategy, it is essential to carefully design the network architecture for its implementation. Current neural field methods [16, 21, 29] commonly employ the rMLP (real-valued MLP) with various nonlinearities as the basic building block. However, they are not suitable for our case due to the inherently complex-valued nature of the wave formulation. Besides, forcibly using two separate real-valued networks to fit the fundamental components of the wave, whether represented as real and imaginary parts or as magnitude and phase, would not only fail to capture the interdependencies among the polarimetric parameters but also significantly increase the parameter count, resulting in a heavy computational burden. Therefore, we design a lightweight complex-valued neural network that naturally aligns with the formulation of the wave, enabling efficient fitting.

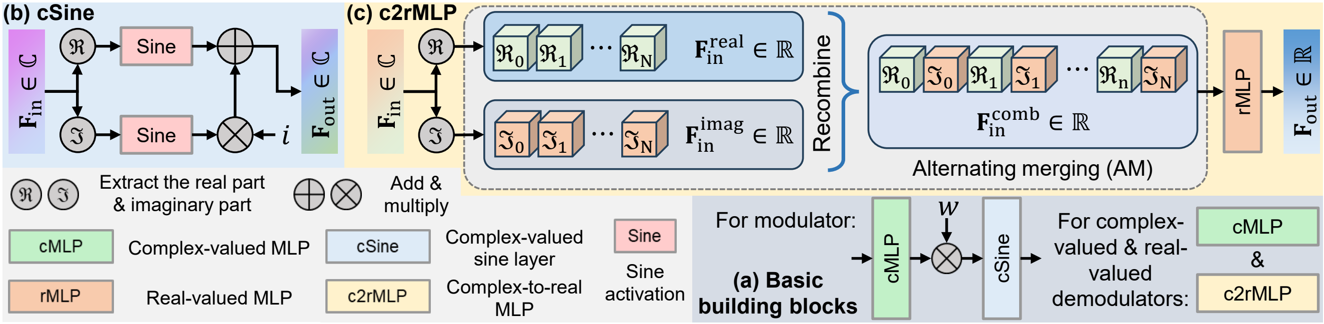

Figure 4. We design a lightweight complex-valued neural network that naturally aligns with the formulation of the wave, enabling efficient fitting. (a) Basic building blocks. (b) The workflow of the proposed cSine layer. (c) The workflow of the proposed c2rMLP layer.

Basic building blocks. As shown in Fig. 4 (a), for the modulators, since they process both complex-valued input and output signals, we design their basic building block as a cMLP (complex-valued MLP) and a custom cSine layer (where the output of cMLP is scaled by a real-valued factor w before being passed to the cSine layer), leveraging the high representational ability of MLPs equipped with periodic activation functions [34]; for the demodulators, the complex-valued ones use a cMLP to decode the complex-valued features into complex-valued output, while the real-valued ones adopt a custom c2rMLP layer to convert the complex-valued features into real-valued output.

cSine layer. As shown in Fig. 4 (b), to ensure its periodicity and boundedness, we design the cSine layer as

$$\mathbf{F}_{\text{out}} = \sin(\Re(\mathbf{F}_{\text{in}})) + i \cdot \sin(\Im(\mathbf{F}_{\text{in}})), \tag{15}$$

where $\Re(\cdot)$ and $\Im(\cdot)$ are the operators for extracting the real and imaginary parts, $\mathbf{F}_{\text{in}}$ and $\mathbf{F}_{\text{out}}$ denote the input and output respectively. The reason we do not apply the sine activation directly to $\mathbf{F}_{\text{in}}$ is that the sine function in the complex domain is not generally bounded (*i.e.*, for $a, b \in \mathbb{R}$, $\sin(a + b \cdot i) = \sin(a)\cosh(b) + \cos(a)\sinh(b)$, where $\cosh(b)$ and $\sinh(b)$ grow exponentially as b increases).

c2rMLP layer. As shown in Fig. 4 (c), the c2rMLP layer is designed to convert the complex-valued input $\mathbf{F}_{\text{in}}$ into real-valued output $\mathbf{F}_{\text{out}}$. Previous works [2, 13] have demonstrated that optimizing values in Cartesian coordinates (*i.e.*, as real and imaginary parts) is less nonlinear and offers higher numerical stability compared to polar coordinates (*i.e.*, as magnitude and phase). Therefore, we first extract the real and imaginary parts of $\mathbf{F}_{\text{in}}$, denoted as $\mathbf{F}^{\text{real}}_{\text{in}}$ and $\mathbf{F}^{\text{imag}}_{\text{in}}$, respectively. Next, we apply an alternating merging (AM) approach to recombine $\mathbf{F}^{\text{real}}_{\text{in}}$ and $\mathbf{F}^{\text{imag}}_{\text{in}}$ into a new feature vector $\mathbf{F}^{\text{comb}}_{\text{in}}$, ensuring that the real and imaginary parts of each element from $\mathbf{F}_{\text{in}}$ are placed at adjacent indices in $\mathbf{F}^{\text{comb}}_{\text{in}}$. Compared to directly concatenating $\mathbf{F}^{\text{real}}_{\text{in}}$ and $\mathbf{F}^{\text{imag}}_{\text{in}}$, the AM approach allows the network to more effectively capture the relationship between the real and imaginary parts. Finally, we use an rMLP to obtain $\mathbf{F}_{\text{out}}$ from $\mathbf{F}^{\text{comb}}_{\text{in}}$. The workflow of the c2rMLP layer can be summarized as:

$$\mathbf{F}_{\text{out}} = rMLP(AM(\Re(\mathbf{F}_{\text{in}}), \Im(\mathbf{F}_{\text{in}}))). \tag{16}$$

Layer configurations. The hidden dimension of the cMLP in all modulators is set to 256. For $\mathcal{F}^{\text{wave}}_{\text{enc}}$, $\mathcal{F}^{\text{msg}}_{\text{enc}}$, $\mathcal{F}^{\text{carr}}_{\text{enc}}$, $\mathcal{F}^{\mathbf{k}}_{\text{enc}}$, and $\mathcal{F}^{\boldsymbol{\theta}}_{\text{enc}}$, we set the number of basic building blocks and the factor w to $(1, 60)$, $(3, 30)$, $(1, 30)$, $(3, 30)$, and $(3, 30)$, respectively. All demodulators consist of a single basic building block. All MLPs are initialized as [34].

Loss function. As shown in Fig. 3, we compute the loss functions between the predicted parameters $\mathbf{I}$, $\mathbf{k}$, and $\boldsymbol{\theta}$ and their corresponding ground truths. The loss function for each parameter is defined as $L_{\mathbf{v}} = \lambda_1 L_1(\mathbf{v}, \mathbf{v}_{\text{gt}}) + \lambda_2 L_2(\mathbf{v}, \mathbf{v}_{\text{gt}})$, where $\mathbf{v}$ and $\mathbf{v}_{\text{gt}}$ denote the predicted parameter and its ground truth, $L_{1,2}$ are the $\ell_{1,2}$ loss functions. For $\mathbf{I}$, $\mathbf{k}$, and $\boldsymbol{\theta}$, we set $\lambda_{1,2}$ to $(2, 20)$, $(0.5, 100)$, and $(1, 10)$, respectively.

Training. Our network is implemented using PyTorch and trained on an NVIDIA 4090 GPU. We train the network for 10^5 epochs with a fixed learning rate of 5×10^{-5}. We use Adam optimizer [11] for optimization.

4. Experiments

4.1. Comparison with Existing Methods

We compare our method with the state-of-the-art multi-band solution (pCON [29]), as well as two state-of-the-art methods based on direct fitting (FINER [21] and S-INR [16]). The evaluation dataset is the same as that used by pCON [29], and we follow its pre-processing approach to resize the polarized images to 1024×1024. Note that for pCON [29], we use the band thresholds specified in its supplementary materials when available; otherwise, we adopt the default values from its released code. The number of parameters and output values for FINER [21] and S-INR [16] are also adjusted following the same approach as in pCON [29], since these methods were originally designed to fit a single quantity at a time, not fitting TI, DoP, and AoP concurrently.

To evaluate the fitting accuracy, following pCON [29], we calculate PSNR and SSIM on both TI, DoP, and AoP. Quan-

Table 1. Quantitative comparisons with the state-of-the-art multi-band solution (pCON [29]), as well as two state-of-the-art methods based on direct fitting (FINER [21], and S-INR [16]). Throughout this paper, ↑ (↓) indicates that higher (lower) values correspond to better results, and we use the **red** (**blue**) text to highlight the best (second-best) results.

	PSNR↑/SSIM↑ of the TI ($\mathbf{I}$)				PSNR↑/SSIM↑ of the DoP ($\mathbf{p}$)				PSNR↑/SSIM↑ of the AoP ($\boldsymbol{\theta}$)			
	pCON [29]	FINER [21]	S-INR [16]	Ours	pCON [29]	FINER [21]	S-INR [16]	Ours	pCON [29]	FINER [21]	S-INR [16]	Ours
Building	**38.64**/**0.942**	33.32/0.830	36.85/0.918	**47.70**/**0.989**	26.84/0.644	**27.44**/0.679	26.63/**0.684**	**29.23**/**0.767**	**21.28**/0.710	19.38/**0.721**	20.81/**0.723**	**21.33**/0.622
Firewood	**39.95**/**0.957**	27.36/0.647	36.75/0.920	**53.54**/**0.992**	**33.66**/0.775	30.61/0.700	33.56/**0.785**	**36.01**/**0.857**	**17.24**/**0.704**	16.78/0.661	17.17/0.653	**17.26**/**0.668**
Grater	**38.62**/**0.955**	27.84/0.732	37.35/0.940	**50.30**/**0.994**	30.13/0.697	26.88/0.543	**30.14**/**0.728**	**32.97**/**0.830**	**17.47**/**0.682**	16.01/0.630	16.48/0.631	**16.92**/**0.637**
Pottery	**39.03**/**0.957**	28.12/0.749	36.18/0.925	**48.54**/**0.986**	32.09/0.769	29.59/0.720	**32.15**/**0.796**	**35.14**/**0.864**	**18.01**/**0.772**	17.36/0.697	17.37/**0.722**	**17.64**/0.677
Stream	**37.58**/**0.945**	34.21/0.879	36.39/0.935	**43.56**/**0.987**	**31.90**/**0.797**	30.37/0.731	30.44/0.730	**33.24**/**0.851**	**24.59**/**0.786**	24.27/0.763	23.97/0.769	**24.65**/**0.788**
Sunroom	**42.94**/**0.968**	35.91/0.910	40.87/0.960	**46.17**/**0.969**	**38.49**/**0.919**	36.41/0.885	38.00/0.909	**41.35**/**0.952**	**28.18**/**0.852**	28.16/0.845	28.13/0.828	**28.20**/**0.857**
Valentines	**35.98**/**0.935**	28.68/0.771	35.84/0.927	**46.60**/**0.985**	30.81/0.750	28.04/0.674	**31.51**/**0.805**	**34.55**/**0.868**	**18.78**/**0.717**	18.04/0.649	18.45/**0.745**	**18.56**/0.618

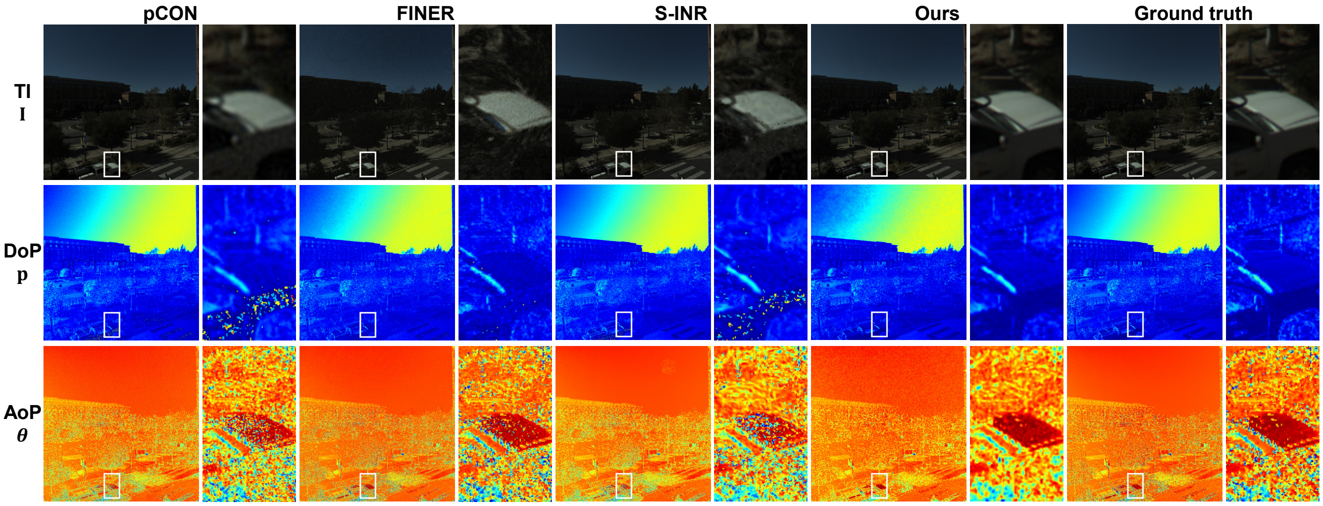

Figure 5. Qualitative comparisons with pCON [29], FINER [21], and S-INR [16]. Please zoom in for better details.

titative results are presented in Tab. 1, where our method significantly outperforms all compared methods on TI and DoP while achieving comparable performance on AoP. Visual quality comparisons are shown in Fig. 5[3]. As illustrated, FINER [21] and S-INR [16] tend to introduce noise patterns in TI, while pCON [29] produces blurry edges; in contrast, our method generates clean and sharp edges. For DoP, all compared methods exhibit severe artifacts, whereas our method does not. For AoP, our method closely resembles the ground truth without distortion. To evaluate the fitting efficiency, we measure both the number of parameters and the model size. The results, presented in Tab. 2, demonstrate that our method outperforms multi-band solutions [29] in fitting accuracy while maintaining similar fitting efficiency to methods based on direct fitting [16, 21].

4.2. Ablation Study

We conduct several ablation studies in Tab. 3 to verify the validity of each design choice, using the scene in Fig. 1 (right) as the validation case. First, we show the importance of representing the polarimetric parameters as a unified wave by comparing to a model that also employs a complex-valued network (composed of the same basic building blocks in Fig. 4 (c)) but instead represents the polarimetric parameters as independent particles (W/o wave). While this model achieves the best performance on AoP, it fails to fit DoP due to the large domain gap among TI, DoP, and AoP. Next, we highlight the significance of our fitting strategy by compar-

Table 2. Comparisons on the number of parameters (Params.) and model size (Sz.) with pCON [29], FINER [21], and S-INR [16]. "Discrete" means directly storing the data in discrete form.

	pCON [29]	FINER [21]	S-INR [16]	Ours	Discrete
Params.↓	1.34M	**0.66M**	**1.27M**	**0.66M**	N/A
Sz.↓	5.14MB	**2.52MB**	5.10MB	**5.05MB**	36.00MB

Table 3. Quantitative evaluation results of ablation study.

	PSNR↑/SSIM↑ of the TI ($\mathbf{I}$)	PSNR↑/SSIM↑ of the DoP ($\mathbf{p}$)	PSNR↑/SSIM↑ of the AoP ($\boldsymbol{\theta}$)
W/o wave	39.22/0.942	6.63/0.020	**18.72**/**0.750**
Direct fitting	39.43/0.969	**39.76**/**0.930**	11.71/0.512
W/o AoP branch	**53.53**/**0.991**	**37.93**/**0.903**	11.99/0.461
Real-valued network	14.42/0.011	31.16/0.746	16.31/0.611
W/o AM	53.51/**0.991**	35.98/0.856	17.04/0.663
Our complete method	**53.54**/**0.992**	36.01/0.857	**17.26**/**0.668**

---

[3] Additional results can be found in the supplementary material.

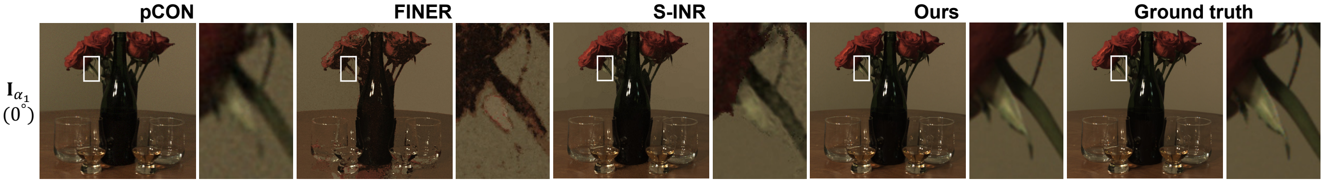

Figure 6. Qualitative comparisons on the accuracy of the rendered polarized images with pCON [29], FINER [21], and S-INR [16]. Here, we show the rendered $\mathbf{I}_{\alpha_1}$ (the polarized images at $\alpha_1 = 0^\circ$). Please zoom in for better details.

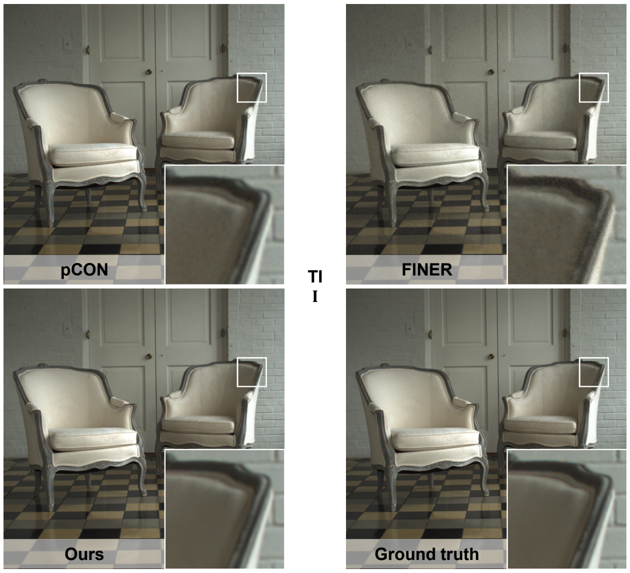

Figure 7. Qualitative comparisons of the 2× super-resolution querying performance with pCON [29] and FINER [21].

ing to a model that directly fits $\mathbf{I}$ and $\mathbf{k}$ of the wave (Direct fitting). While this model achieves the best performance on the DoP, it cannot fit AoP well because it optimizes AoP in an implicit manner, which proves ineffective. Besides, we validate the advantage of introducing a dedicated branch (including the modulator $\mathcal{F}^{\boldsymbol{\theta}}_{\text{enc}}$ and the demodulator $\mathcal{F}^{\boldsymbol{\theta}}_{\text{dec}}$) specifically for AoP by removing it (W/o AoP branch). We can see that incorporating this branch notably enhances AoP performance, with only a minimal trade-off in DoP accuracy. Then, we verify the necessity of our complex-valued network by comparing to a model that forcibly uses two separate real-valued networks (replacing our cMLP and cSine with rMLP and Sine activation) to fit the real and imaginary parts of the wave (Real-valued network). The results show that this model fails to fit TI, as it struggles to capture the interdependencies among the polarimetric parameters. Finally, we also compare with a model that does not adopt the AM approach within the c2rMLP (W/o AM), which shows that AM plays a crucial role in enhancing the performance. The above experiments indicate that our complete method effectively balances the performance of all polarimetric parameters.

4.3. Application

Polarized image rendering. To show that our method can facilitate polarization-based vision, we use the reconstructed polarimetric parameters to render polarized images and evaluate their accuracy. For validation, we select polarizer angles of $\alpha_{1,2,3,4} = 0^\circ, 45^\circ, 90^\circ, 135^\circ$, which are the specific angles that real polarization cameras can capture. Visual quality comparisons are shown in Fig. 6[4], where our result has fewer artifacts and noise patterns.

Super-resolution querying. To demonstrate the continuity of the reconstructed results, we evaluate the super-resolution querying performance. Specifically, we fit the polarimetric parameters at a resolution of 1024×1024 and query them at 2048×2048. At this stage, we compare only with pCON [29] and FINER [21], as S-INR [16] supports querying only at the resolution used for fitting. Visual quality comparisons are shown in Fig. 6[5], where our result exhibits finer details and better fidelity to the ground truth.

5. Conclusion

We propose a polarimetric neural field method to reconstruct continuous polarimetric parameters from discrete data. By introducing a physics-grounded representation scheme that represents the parameters as a unified complex-valued wave, combined with a tuning-free fitting strategy and a lightweight complex-valued neural network, it achieves state-of-the-art performance and facilitates smooth polarized image rendering and flexible resolution adjustments.

Limitations. Since our method is specifically designed for linear polarization, similar to pCON [29], it is not applicable to circular polarization. As future work, we also aim to extend our method to support circular polarization.

Acknowledgment

This work was supported in part by JST-Mirai Program Grant Number JPMJM123G1, National Natural Science Foundation of China (Grant No. 62136001, 62088102, 62472044, U24B20155), Hebei Natural Science Foundation Project No. 242Q0101Z, and Beijing-Tianjin-Hebei Basic Research Funding Program No. F2024502017.

---

[4] Additional results (as well as quantitative evaluations for all selected polarizer angles) can be found in the supplementary material.

[5] Additional results can be found in the supplementary material.

References

[1] Robert B Ash. *Information theory*. Courier Corporation, 2012. 3

[2] Joshua Bassey, Lijun Qian, and Xianfang Li. A survey of complex-valued neural networks. *arXiv preprint arXiv:2101.12249*, 2021. 6

[3] Akshat Dave, Yongyi Zhao, and Ashok Veeraraghavan. PANDORA: Polarization-aided neural decomposition of radiance. In *Proc. of European Conference on Computer Vision*, pages 538–556, 2022. 3

[4] Louis De Broglie. Waves and quanta. *Nature*, 112(2815): 540–540, 1923. 1, 4

[5] Boyang Deng, John P Lewis, Timothy Jeruzalski, Gerard Pons-Moll, Geoffrey Hinton, Mohammad Norouzi, and Andrea Tagliasacchi. Nasa: Neural articulated shape approximation. In *Proc. of European Conference on Computer Vision*, pages 612–628, 2020. 3

[6] Valentin Deschaintre, Yiming Lin, and Abhijeet Ghosh. Deep polarization imaging for 3D shape and SVBRDF acquisition. In *Proc. of Computer Vision and Pattern Recognition*, 2021. 1, 2

[7] Eugene Hecht. *Optics*. Pearson Education India, 2012. 3

[8] Achuta Kadambi, Vage Taamazyan, Boxin Shi, and Ramesh Raskar. Depth sensing using geometrically constrained polarization normals. *International Journal of Computer Vision*, 125:34–51, 2017. 2

[9] Agastya Kalra, Vage Taamazyan, Supreeth Krishna Rao, Kartik Venkataraman, Ramesh Raskar, and Achuta Kadambi. Deep polarization cues for transparent object segmentation. In *Proc. of Computer Vision and Pattern Recognition*, pages 8602–8611, 2020. 2, 3

[10] Youngchan Kim, Wonjoon Jin, Sunghyun Cho, and Seung-Hwan Baek. Neural spectro-polarimetric fields. In *Proc. of ACM SIGGRAPH Asia*, pages 1–11, 2023. 3

[11] Diederik P Kingma and Jimmy Ba. ADAM: A method for stochastic optimization, 2014. 6

[12] GP Können. *Polarized light in nature*. CUP Archive, 1985. 1, 3

[13] ChiYan Lee, Hideyuki Hasegawa, and Shangce Gao. Complex-valued neural networks: A comprehensive survey. *IEEE/CAA Journal of Automatica Sinica*, 9(8):1406–1426, 2022. 6

[14] Chenyang Lei, Xuhua Huang, Mengdi Zhang, Qiong Yan, Wenxiu Sun, and Qifeng Chen. Polarized reflection removal with perfect alignment in the wild. In *Proc. of Computer Vision and Pattern Recognition*, pages 1750–1758, 2020. 1, 2

[15] Chenhao Li, Taishi Ono, Takeshi Uemori, Hajime Mihara, Alexander Gatto, Hajime Nagahara, and Yusuke Moriuchi. NeISF: Neural incident stokes field for geometry and material estimation. In *Proc. of Computer Vision and Pattern Recognition*, pages 21434–21445, 2024. 3

[16] Jiayi Li, Xile Zhao, Jianli Wang, Chao Wang, and Min Wang. Superpixel-informed implicit neural representation for multi-dimensional data. In *Proc. of European Conference on Computer Vision*, pages 258–276, 2024. 1, 3, 4, 5, 6, 7, 8

[17] Jason Chun Lok Li, Chang Liu, Binxiao Huang, and Ngai Wong. Learning spatially collaged fourier bases for implicit neural representation. In *Proc. of the AAAI Conference on Artificial Intelligence*, pages 13492–13499, 2024. 3

[18] Ning Li, Yongqiang Zhao, Quan Pan, Seong G Kong, and Jonathan Cheung-Wai Chan. Full-time monocular road detection using zero-distribution prior of angle of polarization. In *Proc. of European Conference on Computer Vision*, pages 457–473, 2020. 2

[19] Zhuoxiao Li, Haiyang Jiang, Mingdeng Cao, and Yinqiang Zheng. Polarized color image denoising. In *Proc. of Computer Vision and Pattern Recognition*, pages 9873–9882, 2023. 2

[20] Yupeng Liang, Ryosuke Wakaki, Shohei Nobuhara, and Ko Nishino. Multimodal material segmentation. In *Proc. of Computer Vision and Pattern Recognition*, pages 19800–19808, 2022. 2

[21] Zhen Liu, Hao Zhu, Qi Zhang, Jingde Fu, Weibing Deng, Zhan Ma, Yanwen Guo, and Xun Cao. FINER: Flexible spectral-bias tuning in implicit neural representation by variable-periodic activation functions. In *Proc. of Computer Vision and Pattern Recognition*, pages 2713–2722, 2024. 1, 2, 3, 4, 5, 6, 7, 8

[22] Youwei Lyu, Zhaopeng Cui, Si Li, Marc Pollefeys, and Boxin Shi. Physics-guided reflection separation from a pair of unpolarized and polarized images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2):2151–2165, 2022. 1, 2

[23] Youwei Lyu, Lingran Zhao, Si Li, and Boxin Shi. Shape from polarization with distant lighting estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(11): 13991–14004, 2023. 1, 2

[24] Julien N. P. Martel, David B. Lindell, Connor Z. Lin, Eric R. Chan, Marco Monteiro, and Gordon Wetzstein. ACORN: Adaptive coordinate networks for neural scene representation. *ACM Transactions on Graphics (Proc. of ACM SIGGRAPH)*, 40(4), 2021. 3

[25] Ricardo Martin-Brualla, Noha Radwan, Mehdi SM Sajjadi, Jonathan T Barron, Alexey Dosovitskiy, and Daniel Duckworth. NeRF in the wild: Neural radiance fields for unconstrained photo collections. In *Proc. of Computer Vision and Pattern Recognition*, pages 7210–7219, 2021. 3

[26] Haiyang Mei, Bo Dong, Wen Dong, Jiaxi Yang, Seung-Hwan Baek, Felix Heide, Pieter Peers, Xiaopeng Wei, and Xin Yang. Glass segmentation using intensity and spectral polarization cues. In *Proc. of Computer Vision and Pattern Recognition*, pages 12622–12631, 2022. 2

[27] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *Proc. of European Conference on Computer Vision*, pages 405–421, 2020. 3

[28] Jeong Joon Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. DeepSDF: Learning continuous signed distance functions for shape representation. In *Proc. of Computer Vision and Pattern Recognition*, pages 165–174, 2019. 3

[29] Henry Peters, Yunhao Ba, and Achuta Kadambi. pCON: Polarimetric coordinate networks for neural scene representations. In *Proc. of Computer Vision and Pattern Recognition*, pages 16579–16589, 2023. 1, 2, 3, 4, 5, 6, 7, 8

[30] Mara Pistellato, Filippo Bergamasco, Tehreem Fatima, and Andrea Torsello. Deep demosaicing for polarimetric filter array cameras. *IEEE Transactions on Image Processing*, 31: 2017–2026, 2022. 2

[31] Sameera Ramasinghe and Simon Lucey. Beyond periodicity: Towards a unifying framework for activations in coordinate-MLPs. In *Proc. of European Conference on Computer Vision*, pages 142–158, 2022. 3

[32] Vishwanath Saragadam, Daniel LeJeune, Jasper Tan, Guha Balakrishnan, Ashok Veeraraghavan, and Richard G Baraniuk. WIRE: Wavelet implicit neural representations. In *Proc. of Computer Vision and Pattern Recognition*, pages 18507–18516, 2023. 3

[33] Yoav Y Schechner, Srinivasa G Narasimhan, and Shree K Nayar. Instant dehazing of images using polarization. In *Proc. of Computer Vision and Pattern Recognition*, pages I–I, 2001. 2, 3

[34] Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. In *Proc. of Advances in Neural Information Processing Systems*, pages 7462–7473, 2020. 3, 6

[35] Yu Sun, Jiaming Liu, Mingyang Xie, Brendt Wohlberg, and Ulugbek S Kamilov. CoIL: Coordinate-based internal learning for tomographic imaging. *IEEE Transactions on Computational Imaging*, 7:1400–1412, 2021. 3

[36] Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. In *Proc. of Advances in Neural Information Processing Systems*, pages 7537–7547, 2020. 2

[37] Chaoran Tian, Weihong Pan, Zimo Wang, Mao Mao, Guofeng Zhang, Hujun Bao, Ping Tan, and Zhaopeng Cui. DPS-Net: Deep polarimetric stereo depth estimation. In *Proc. of International Conference on Computer Vision*, pages 3569–3579, 2023. 2

[38] Yuehao Wang, Yonghao Long, Siu Hin Fan, and Qi Dou. Neural rendering for stereo 3D reconstruction of deformable tissues in robotic surgery. In *Proc. of International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 431–441, 2022. 3

[39] Yiheng Xie, Towaki Takikawa, Shunsuke Saito, Or Litany, Shiqin Yan, Numair Khan, Federico Tombari, James Tompkin, Vincent Sitzmann, and Srinath Sridhar. Neural fields in visual computing and beyond. In *Computer Graphics Forum*, pages 641–676, 2022. 2

[40] Chu Zhou, Minggui Teng, Yufei Han, Chao Xu, and Boxin Shi. Learning to dehaze with polarization. In *Proc. of Advances in Neural Information Processing Systems*, pages 11487–11500, 2021. 2

[41] Chu Zhou, Yufei Han, Minggui Teng, Jin Han, Si Li, Chao Xu, and Boxin Shi. Polarization guided HDR reconstruction via pixel-wise depolarization. *IEEE Transactions on Image Processing*, 32:1774–1787, 2023. 2

[42] Chu Zhou, Minggui Teng, Youwei Lyu, Si Li, Chao Xu, and Boxin Shi. Polarization-aware low-light image enhancement. In *Proc. of the AAAI Conference on Artificial Intelligence*, pages 3742–3750, 2023. 2, 3

[43] Chu Zhou, Yixing Liu, Chao Xu, and Boxin Shi. Quality-improved and property-preserved polarimetric imaging via complementarily fusing. In *Proc. of Advances in Neural Information Processing Systems*, 2024. 1, 2, 5

[44] Chu Zhou, Minggui Teng, Xinyu Zhou, Chao Xu, Imari Sato, and Boxin Shi. Learning to deblur polarized images. *International Journal of Computer Vision*, 2025. 2

[45] Chu Zhou, Chao Xu, and Boxin Shi. Polarization guided mask-free shadow removal. In *Proc. of the AAAI Conference on Artificial Intelligence*, pages 10716–10724, 2025. 2

[46] Shuangfan Zhou, Chu Zhou, Youwei Lyu, Heng Guo, Zhanyu Ma, Boxin Shi, and Imari Sato. PIDSR: Complementary polarized image demosaicing and super-resolution. In *Proc. of Computer Vision and Pattern Recognition*, pages 16081–16090, 2025. 2

Supplementary Material:
Polarimetric Neural Field via Unified Complex-Valued Wave Representation

Chu Zhou[1] Yixin Yang[2,3] Junda Liao[1,4] Heng Guo[5] Boxin Shi[2,3*] Imari Sato[1,4*]
[1]National Institute of Informatics, Japan
[2]State Key Laboratory of Multimedia Information Processing, School of Computer Science, Peking University, China
[3]National Engineering Research Center of Visual Technology, School of Computer Science, Peking University, China
[4]Graduate School of Information Science and Technology, University of Tokyo, Japan
[5]School of Artificial Intelligence, Beijing University of Posts and Telecommunications, China
zhou_chu@hotmail.com, {yangyixin93, shiboxin}@pku.edu.cn,
liao-junda@g.ecc.u-tokyo.ac.jp, guoheng@bupt.edu.cn, imarik@nii.ac.jp

A. Discussions About the Conventions

In this section, we provide discussions about the conventions for polarimetric parameters, corresponding to Footnote 2 of the main paper.

Malus' law [1] describes the relationship between the polarized image $\mathbf{I}_\alpha$ (captured using a linear polarizer at angle α) and the polarimetric parameters (including the total intensity (TI) $\mathbf{I}$, degree of polarization (DoP) $\mathbf{p}$, and angle of polarization (AoP) $\boldsymbol{\theta}$). Depending on the convention [9], it can be expressed as:

$$\mathbf{I}_\alpha = \frac{\mathbf{I}}{2} \cdot (1 - \mathbf{p} \cdot \cos(2(\alpha - \boldsymbol{\theta}))), \tag{A}$$

which is adopted in our work and in previous studies [7, 8] (denoted as **Convention1**), or as

$$\mathbf{I}_\alpha = \frac{\mathbf{I}}{2} \cdot (1 + \mathbf{p} \cdot \cos(2(\boldsymbol{\theta} - \alpha))), \tag{B}$$

which is used by pCON [5] (denoted as **Convention2**). Besides, the Stokes parameters [2] also have two different definitions. Denoting $\mathbf{I}_{\alpha_{1,2,3,4}}$ as the polarized images captured at $\alpha_{1,2,3,4} = 0°, 45°, 90°, 135°$, the Stokes parameters $\mathbf{S}_{0,1,2}$ are defined as

$$\begin{cases} \mathbf{S}_0 = \mathbf{I}_{\alpha_1} + \mathbf{I}_{\alpha_3} = \mathbf{I}_{\alpha_2} + \mathbf{I}_{\alpha_4} \\ \mathbf{S}_1 = \mathbf{I}_{\alpha_3} - \mathbf{I}_{\alpha_1}, \quad \text{and} \quad \mathbf{S}_2 = \mathbf{I}_{\alpha_4} - \mathbf{I}_{\alpha_2} \end{cases} \tag{C}$$

if we follow Convention1, and defined as

$$\begin{cases} \mathbf{S}_0 = \mathbf{I}_{\alpha_1} + \mathbf{I}_{\alpha_3} = \mathbf{I}_{\alpha_2} + \mathbf{I}_{\alpha_4} \\ \mathbf{S}_1 = \mathbf{I}_{\alpha_1} - \mathbf{I}_{\alpha_3}, \quad \text{and} \quad \mathbf{S}_2 = \mathbf{I}_{\alpha_2} - \mathbf{I}_{\alpha_4} \end{cases} \tag{D}$$

if we follow Convention2. Since in polarimetric imaging, polarimetric parameters are typically estimated as

$$\mathbf{I} = \mathbf{S}_0, \mathbf{p} = \frac{\sqrt{\mathbf{S}_1^2 + \mathbf{S}_2^2}}{\mathbf{S}_0} \quad \text{and} \quad \boldsymbol{\theta} = \frac{1}{2}\arctan(\frac{\mathbf{S}_2}{\mathbf{S}_1}), \tag{E}$$

the estimated TI $\mathbf{I}$ and DoP $\mathbf{p}$ remain the same under both Convention1 and Convention2. However, the estimated AoP $\boldsymbol{\theta}$ would be different due to the sensitivity of $\arctan$ to small numerical errors (*i.e.*, despite the arguments being mathematically equivalent ($\frac{\mathbf{I}_{\alpha_4} - \mathbf{I}_{\alpha_2}}{\mathbf{I}_{\alpha_3} - \mathbf{I}_{\alpha_1}} = \frac{\mathbf{I}_{\alpha_2} - \mathbf{I}_{\alpha_4}}{\mathbf{I}_{\alpha_1} - \mathbf{I}_{\alpha_3}}$), the limited precision of floating-point arithmetic and the bounded range of $\arctan$ can lead to small discrepancies, especially when dealing with values close to the boundaries).

Here, we also provide some qualitative comparisons with pCON [5], which follows Convention2 (the convention used by pCON [5]). Note that both our method and pCON [5] are retrained using polarimetric parameters according to Convention2. Besides, we also visualize the DoP and AoP directly, using the same visualization approach employed by pCON [5], instead of using color maps as in our main paper. Results are shown in Fig. A, where we can see that our method produces fewer artifacts and noise patterns compared with pCON [5].

B. Additional Qualitative Comparisons

In this section, we provide additional qualitative comparisons with pCON [5], FINER [4], and S-INR [3], corresponding to Footnote 3 of the main paper. The results are shown in Fig. B and Fig. C.

Besides, since the polarized images provided by pCON [5] are captured by polarization cameras, to evaluate the generalization ability of our method, we also provide qualitative comparisons using the polarized images captured via

---

[*] Corresponding authors.

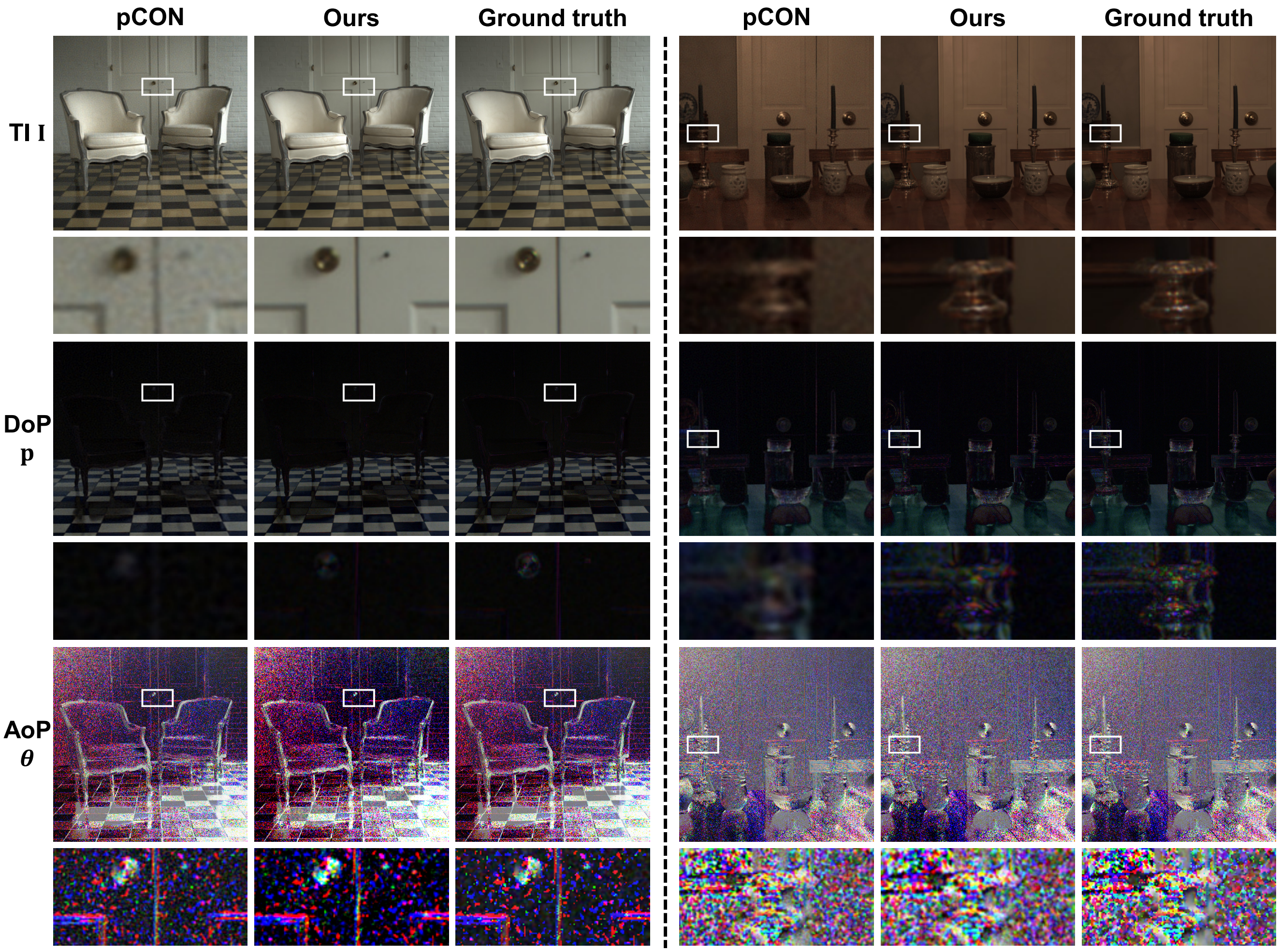

Figure A. Qualitative comparisons with pCON [5] following the convention used by pCON [5] (Convention2). The DoP and AoP are visualized directly, following the same visualization approach employed by pCON [5], instead of using color maps as in our main paper. Please zoom in for better details.

rotating a polarizer [6]. The results are shown in Fig. D and Fig. E.

C. Additional Results of Polarized Image Rendering

In this section, we provide additional qualitative comparisons on the accuracy of the rendered polarized images with pCON [5], FINER [4], and S-INR [3], corresponding to Footnote 4 of the main paper. The results are shown in Fig. F (the rendered polarized images at $\alpha_1 = 0°$) and Fig. G (the rendered polarized images at $\alpha_2 = 45°$).

Besides, we provide the quantitative evaluations for all selected polarizer angles ($\alpha_{1,2,3,4} = 0°, 45°, 90°, 135°$) in Tab. A. From the results, we can see that our method outperforms all compared methods for all polarizer angles.

D. Additional Results of Super-Resolution Querying

In this section, we provide additional qualitative comparisons of the super-resolution querying performance with pCON [5] and FINER [4], corresponding to Footnote 5 of the main paper. Specifically, we fit the polarimetric parameters at a resolution of 1024×1024 and query them at 2048×2048. Note that here we cannot compare with S-INR [3], since it supports querying only at the resolution used for fitting. The results are shown in Fig. H.

References

[1] Eugene Hecht. *Optics*. Pearson Education India, 2012. 1

[2] GP Können. *Polarized light in nature*. CUP Archive, 1985. 1

[3] Jiayi Li, Xile Zhao, Jianli Wang, Chao Wang, and Min Wang. Superpixel-informed implicit neural representation for multi-

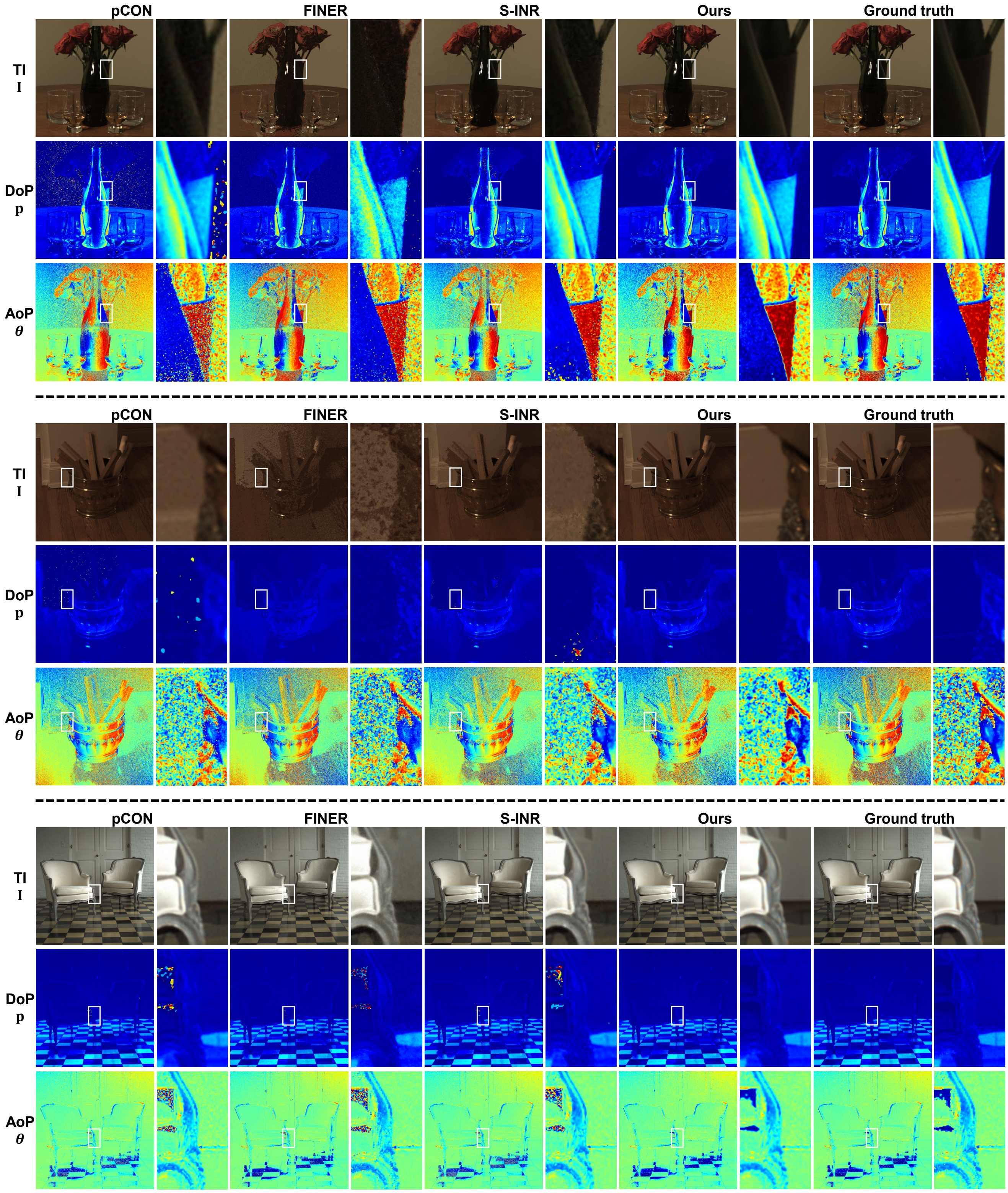

Figure B. Additional qualitative comparisons with pCON [5], FINER [4], and S-INR [3] (part 1). Please zoom in for better details.

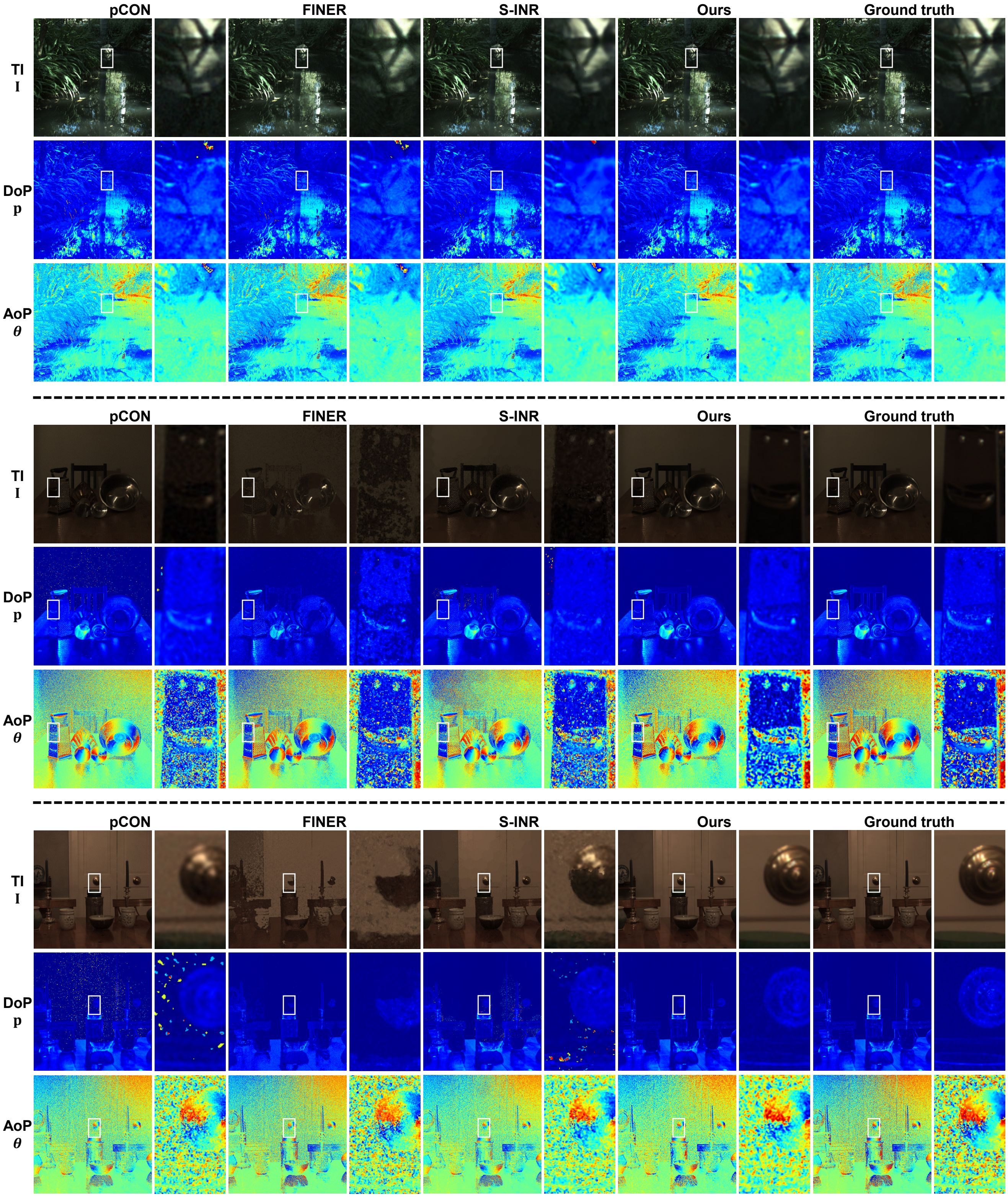

Figure C. Additional qualitative comparisons with pCON [5], FINER [4], and S-INR [3] (part 2). Please zoom in for better details.

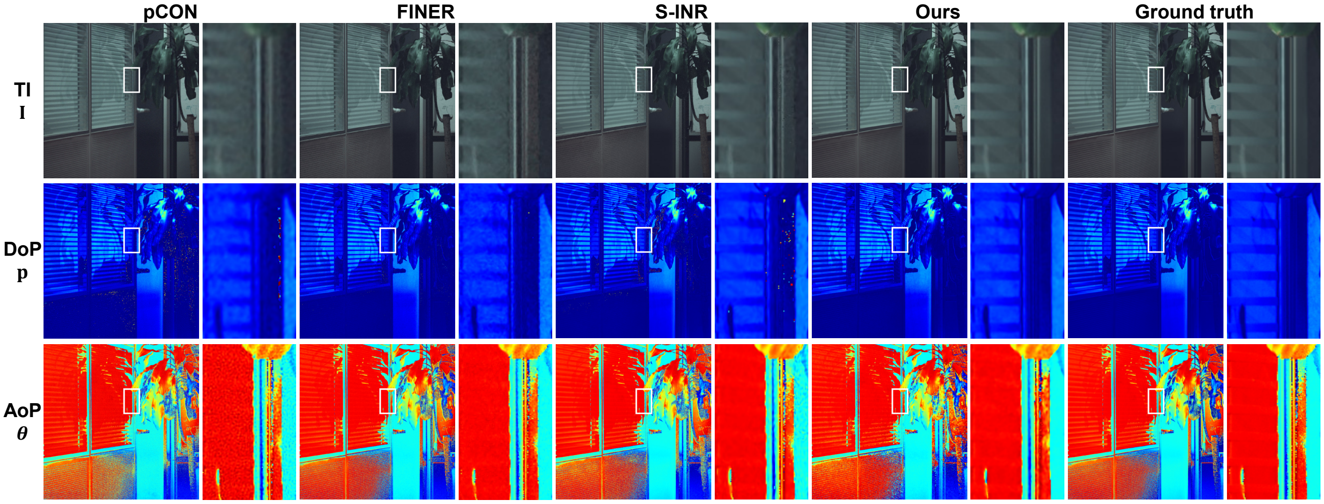

Figure D. Qualitative comparisons with pCON [5], FINER [4], and S-INR [3] using the polarized images captured via rotating a polarizer [6] (part 1). Please zoom in for better details.

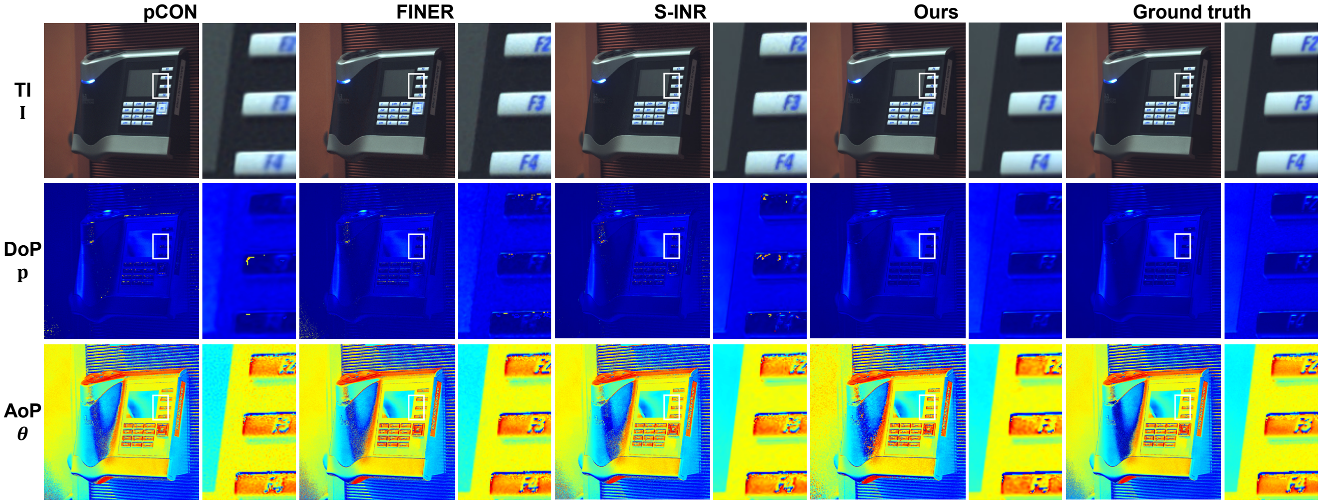

Figure E. Qualitative comparisons with pCON [5], FINER [4], and S-INR [3] using the polarized images captured via rotating a polarizer [6] (part 2). Please zoom in for better details.

dimensional data. In *Proc. of European Conference on Computer Vision*, pages 258–276, 2024. 1, 2, 3, 4, 5, 6

[4] Zhen Liu, Hao Zhu, Qi Zhang, Jingde Fu, Weibing Deng, Zhan Ma, Yanwen Guo, and Xun Cao. FINER: Flexible spectral-bias tuning in implicit neural representation by variable-periodic activation functions. In *Proc. of Computer Vision and Pattern Recognition*, pages 2713–2722, 2024. 1, 2, 3, 4, 5, 6, 7

[5] Henry Peters, Yunhao Ba, and Achuta Kadambi. pCON: Polarimetric coordinate networks for neural scene representations. In *Proc. of Computer Vision and Pattern Recognition*, pages 16579–16589, 2023. 1, 2, 3, 4, 5, 6, 7

[6] Simeng Qiu, Qiang Fu, Congli Wang, and Wolfgang Heidrich. Linear polarization demosaicking for monochrome and colour polarization focal plane arrays. In *Computer Graphics Forum*, pages 77–89, 2021. 2, 5

[7] Yoav Y Schechner, Srinivasa G Narasimhan, and Shree K Nayar. Instant dehazing of images using polarization. In *Proc. of Computer Vision and Pattern Recognition*, pages I–I, 2001. 1

[8] Chu Zhou, Minggui Teng, Youwei Lyu, Si Li, Chao Xu, and Boxin Shi. Polarization-aware low-light image enhancement. In *Proc. of the AAAI Conference on Artificial Intelligence*, pages 3742–3750, 2023. 1

[9] Evgenij Zubko and Ekaterina Chornaya. On the ambiguous definition of the degree of linear polarization. *Research Notes of the AAS*, 3(3):45, 2019. 1

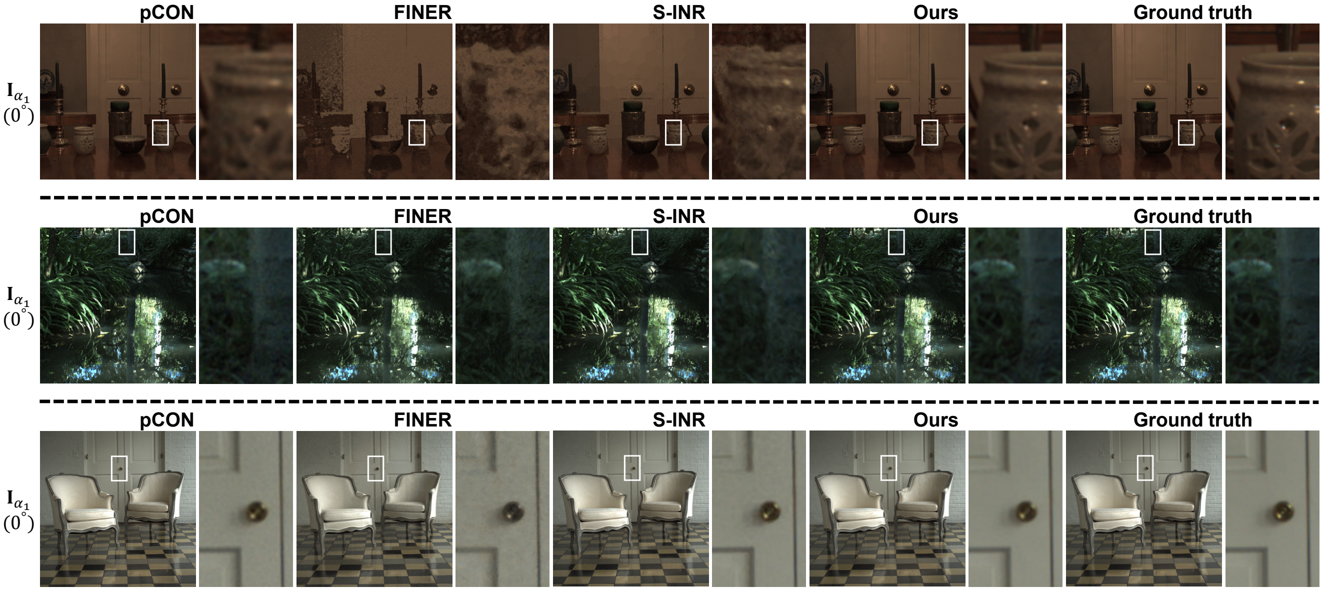

Figure F. Additional qualitative comparisons on the accuracy of the rendered polarized images with pCON [5], FINER [4], and S-INR [3] (part 1). Here, we show the rendered $\mathbf{I}_{\alpha_1}$ (the polarized images at $\alpha_1 = 0°$). Please zoom in for better details.

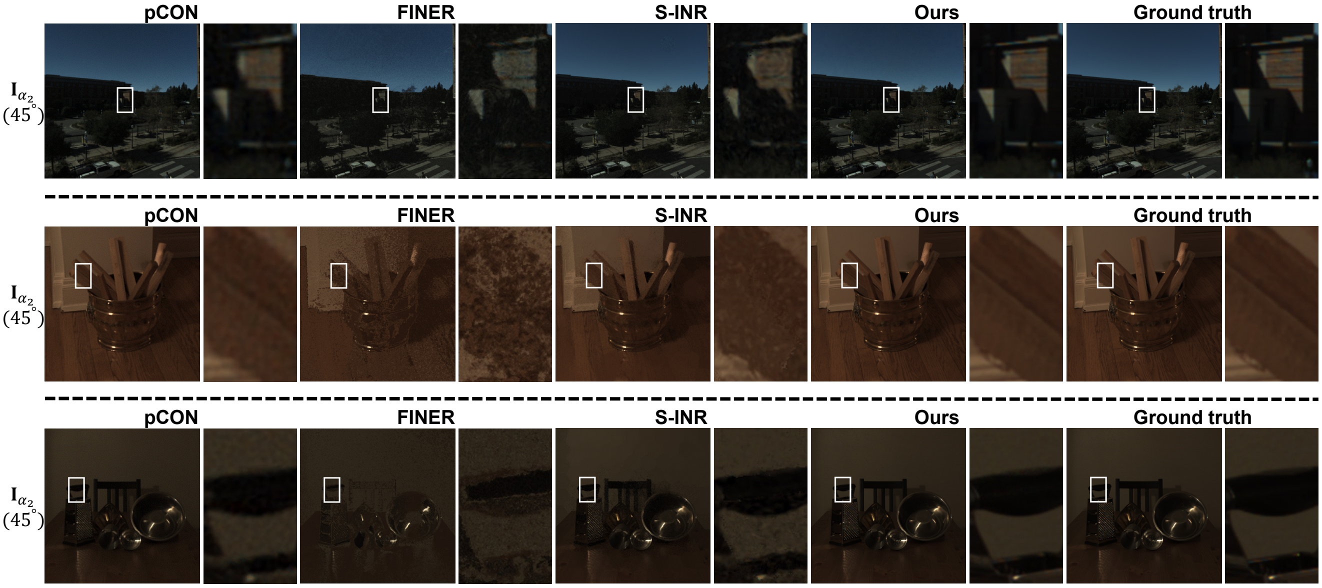

Figure G. Additional qualitative comparisons on the accuracy of the rendered polarized images with pCON [5], FINER [4], and S-INR [3] (part 2). Here, we show the rendered $\mathbf{I}_{\alpha_2}$ (the polarized images at $\alpha_1 = 45°$). Please zoom in for better details.

Table A. Quantitative comparisons on the accuracy of the rendered polarized images with pCON [5], FINER [4], and S-INR [3].

	PSNR↑/SSIM↑ of $\mathbf{I}_{\alpha_1}$ (0°)				PSNR↑/SSIM↑ of $\mathbf{I}_{\alpha_2}$ (45°)				PSNR↑/SSIM↑ of $\mathbf{I}_{\alpha_3}$ (90°)				PSNR↑/SSIM↑ of $\mathbf{I}_{\alpha_4}$ (135°)			
	pCON [5]	FINER [4]	S-INR [3]	Ours	pCON [5]	FINER [4]	S-INR [3]	Ours	pCON [5]	FINER [4]	S-INR [3]	Ours	pCON [5]	FINER [4]	S-INR [3]	Ours
Building	**36.58**/0.884	33.24/0.820	36.45/**0.909**	**39.11**/**0.931**	**36.72**/**0.910**	31.99/0.784	35.65/0.900	**42.78**/**0.971**	**36.27**/0.881	32.81/0.810	36.16/**0.907**	**39.04**/**0.932**	**38.61**/**0.932**	34.23/0.856	37.24/0.921	**43.67**/**0.974**
Firewood	**39.46**/**0.949**	27.22/0.641	36.59/0.914	**47.60**/**0.989**	**39.55**/**0.949**	27.36/0.645	36.58/0.914	**47.74**/**0.989**	**39.52**/**0.949**	27.41/0.647	36.49/0.912	**47.50**/**0.989**	**39.55**/**0.949**	27.32/0.645	36.59/0.914	**47.75**/**0.989**
Grater	**37.92**/**0.946**	27.72/0.729	36.85/0.933	**44.72**/**0.987**	**37.91**/**0.946**	27.76/0.726	36.89/0.934	**45.01**/**0.986**	**38.01**/**0.948**	27.77/0.725	36.81/0.933	**44.88**/**0.987**	**38.19**/**0.946**	27.78/0.728	37.04/0.934	**45.20**/**0.986**
Pottery	**38.04**/**0.947**	27.82/0.734	35.62/0.915	**44.41**/**0.985**	**38.33**/**0.946**	28.08/0.742	35.87/0.918	**44.42**/**0.981**	**38.55**/**0.949**	28.25/0.753	35.87/0.917	**44.90**/**0.985**	**38.34**/**0.945**	28.04/0.742	35.85/0.917	**44.44**/**0.981**
Stream	**35.85**/**0.936**	32.80/0.867	34.88/0.922	**39.88**/**0.981**	**37.18**/**0.939**	34.21/0.879	36.37/0.931	**40.81**/**0.981**	**37.61**/**0.938**	34.71/0.879	36.66/0.933	**42.26**/**0.983**	**35.88**/**0.932**	32.93/0.864	35.03/0.924	**39.51**/**0.980**
Sunroom	**41.22**/**0.966**	35.14/0.891	39.76/0.959	**44.16**/**0.986**	**41.91**/**0.970**	35.71/0.903	40.21/0.963	**44.44**/**0.987**	**41.68**/**0.969**	35.85/0.908	40.06/0.962	**44.32**/**0.987**	**41.79**/**0.970**	35.59/0.902	40.07/0.963	**44.33**/**0.987**
Valentines	34.98/**0.919**	28.44/0.767	**35.26**/0.918	**41.70**/**0.979**	35.14/0.914	28.48/0.757	**35.32**/**0.918**	**40.82**/**0.969**	**35.36**/**0.924**	28.47/0.759	35.23/0.917	**41.85**/**0.980**	35.11/0.914	28.46/0.761	**35.34**/**0.918**	**40.82**/**0.969**

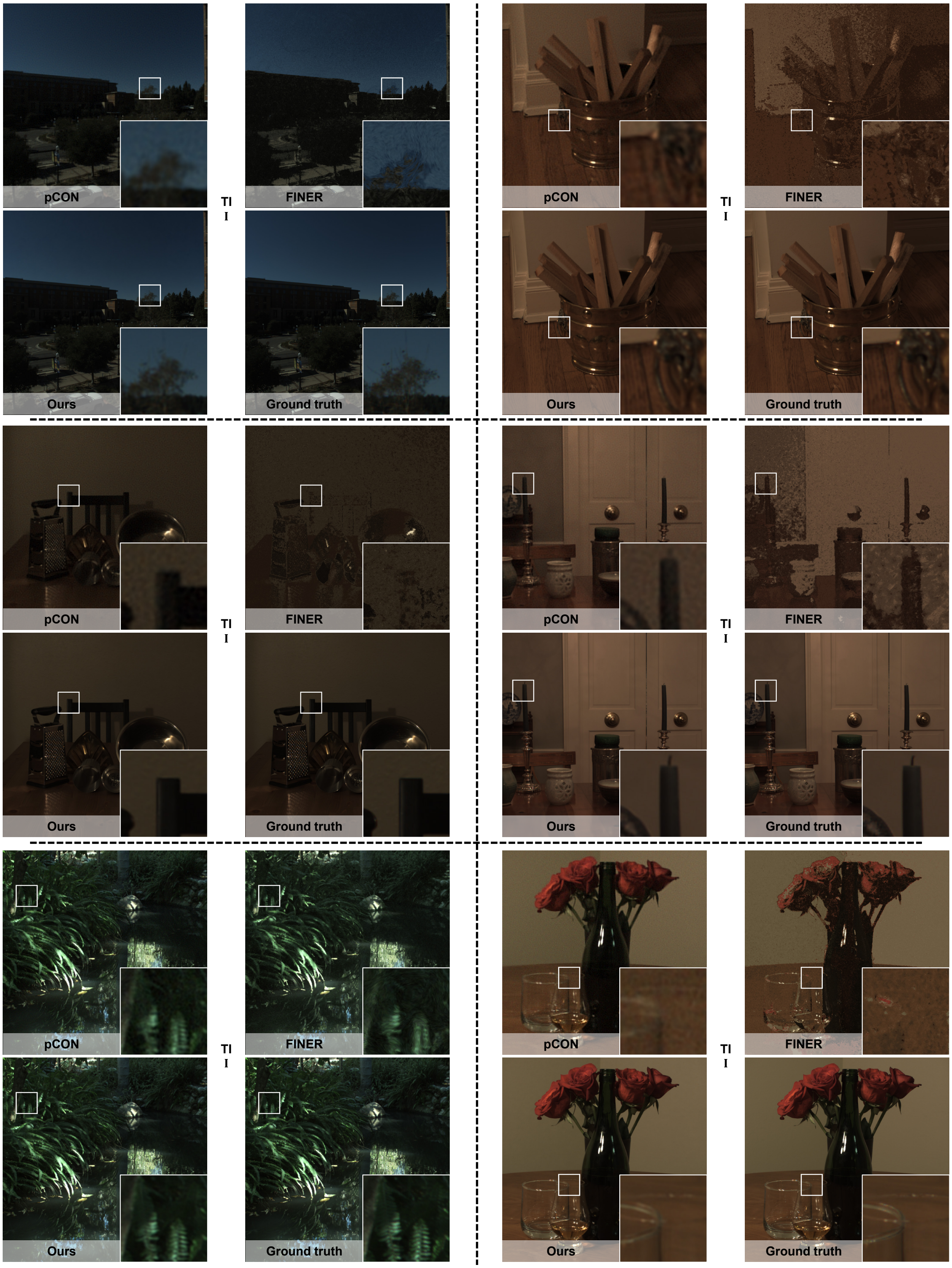

Figure H. Additional qualitative comparisons of the 2× super-resolution querying performance with pCON [5] and FINER [4].