pCON: Polarimetric Coordinate Networks for Neural Scene Representations

Henry Peters[*,1], Yunhao Ba[*,2], Achuta Kadambi[1,2]
[1]Computer Science Department, University of California, Los Angeles (UCLA)
[2]Electrical and Computer Engineering Department, UCLA
hpeters@ucla.edu, yhba@ucla.edu, achuta@ee.ucla.edu

Supplementary contents

This supplement is organized as follows:

- Section A presents results for training models on RGB polarized images, as opposed to training them on $\mathbf{\Phi}$, $\boldsymbol{\rho}$ and $\mathbf{I}_{un}$ maps directly.
- Section B presents results for training baseline models on $\mathbf{\Phi}$ maps only.
- Section C presents a comparison between our full model and baselines trained with twice as many parameters as their originally proposed models.
- Section D provides a description of the code and dataset that we upload along with this supplement.
- Section E provides links to the training code for all baseline models.
- Section F provides more details on our implementation for reproducibility.

A. Training on RGB domain

In the main paper, we trained all models directly to predict each of the three polar quantities of interest. Here, we present an alternative solution by training our model and all baselines on the four polarization images directly. We train each model to predict the RGB value at each coordinate for images taken of a scene through filters at 0, 45, 90 and 135 degrees. We then calculate $\mathbf{\Phi}$, $\boldsymbol{\rho}$ and $\mathbf{I}_{un}$ maps from the predictions of each polar quantity using the math described in Sec. 3.1. Qualitative and quantitative results can be found in Fig. A. All models achieve similar performance on $\mathbf{I}_{un}$ and $\boldsymbol{\rho}$ maps, but ours still achieves a higher SSIM on the predicted $\mathbf{\Phi}$ map. Under this arrangement, the performance of all models on AoLP is poor, which means the underlying polarization information is not preserved properly. This is why we decided to train directly on the maps for each polar quantity, rather than the RGB polarized images in the main paper. It is worth noting that our model still achieves the highest SSIM on $\mathbf{\Phi}$ of all networks considered under this configuration.

B. Learning AoLP individually

In the main paper, we trained all baseline models to predict AoLP ($\mathbf{\Phi}$), DoLP ($\boldsymbol{\rho}$) and unpolarized intensity ($\mathbf{I}_{un}$) concurrently. To further demonstrate our model's improvement over the baselines, we present results for training baseline models to predict only $\mathbf{\Phi}$ maps. For the comparison models in this section we use roughly one third of the parameters used in the models trained to predict all three quantities at once. Our jointly trained model still outperforms all baselines in terms of SSIM. See Fig. B for the visualization of this experiment.

C. Baselines with double the number of parameters

Our full model used more parameters than the baselines we compared to. In the main paper, we presented results on both our full model and our half model, which had the same number of parameters as the baselines. Here we instead increase the number of parameters of all the baselines to match the number in our full model. While performance does increase for AoLP ($\mathbf{\Phi}$) maps, all baselines still create artifacts in the DoLP ($\boldsymbol{\rho}$) and $\mathbf{I}_{un}$ maps, and our model still achieves at least as good performance as all baselines when predicting $\mathbf{\Phi}$. Similarly our half model achieves equal or better performance on all three quantities. Our half and full models can both be obtained with a single training, while all baselines require a complete retraining in order to change the number of parameters. A visualization of the results from this experiment can be seen in Fig. C. See Table A for purely quantitative results.

D. Our code and dataset

We provide links to the code and dataset used for this paper on the project webpage. The webpage can be found

[*]Equal contribution.

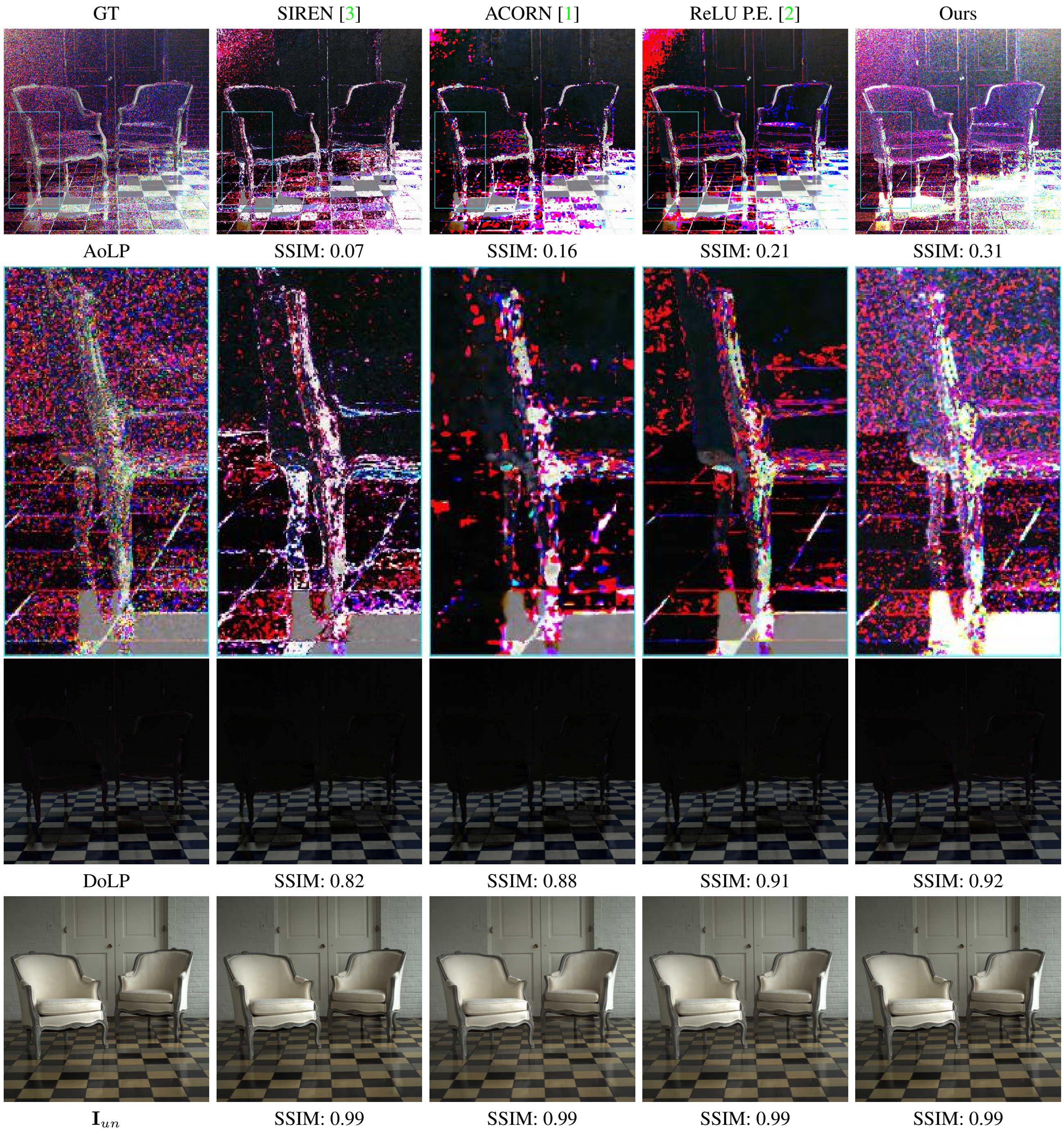

Figure A. **Even though all models predict AoLP poorly when trained on polarization RGB images, our model still achieves the highest SSIM.** Our predicted $\mathbf{\Phi}$ map also looks qualitatively most similar to the baseline. All models in this figure were trained to reconstruct RGB images of a scene taken through polarizing filters at 0, 45, 90 and 135 degrees directly.

here. https://visual.ee.ucla.edu/pcon.htm/

E. Code links for baseline models

For all results presented in the main paper and this supplement, we trained baseline models using code from the GitHub repositories listed in Table B.

GT | SIREN [3] | ACORN [1] | ReLU P.E. [2] | Ours

AoLP | SSIM: 0.69 | SSIM: 0.49 | SSIM: 0.63 | SSIM: 0.77

Figure B. **Even when baseline models are trained on AoLP ($\mathbf{\Phi}$) alone, they still do not outperform our model, which was trained to predict all three polar quantities at once.**

Model	SSIM $\mathbf{\Phi}$	SSIM $\boldsymbol{\rho}$	SSIM $\mathbf{I}_{un}$	# Parameters
SIREN [3]	0.60	0.73	0.59	600K
SIREN [3]	0.77	0.77	0.68	1.3M
ACORN [1]	0.51	0.80	0.77	500K
ACORN [1]	0.62	0.76	0.82	1.6M
ReLU w/P.E. [2]	0.63	0.82	0.71	600K
ReLU w/P.E. [2]	0.68	0.83	0.75	1.3M
Ours (half)	0.62	0.82	0.89	600K
Ours (full)	0.77	0.82	0.89	1.3M

Table A. **Even with twice as many parameters, our model still achieves better SSIM than all baselines.**

Methods	Links
SIREN [3]	https://github.com/vsitzmann/siren
ACORN [1]	https://github.com/computational-imaging/ACORN
ReLU w/P.E. [2]	https://github.com/computational-imaging/ACORN

Table B. **Code links for the comparison methods.**

F. Hyperparameter details

We present the thresholds that we used to allocate singular values to each band of the network for all the scenes that appear in the main paper. We present our values for AoLP ($\mathbf{\Phi}$), DoLP ($\boldsymbol{\rho}$) and $\mathbf{I}_{un}$. These thresholds were hyperparameters that we chose based on the MSE of the image reconstructions using all singular values up to the current threshold.

- Firewood (Fig. 1)
  - AoLP:
    $[180, 360, 540, 720, 900, 920, 940, 960, 980, 1024]$
  - DoLP:
    $[40, 80, 120, 140, 160, 200, 300, 400, 600, 1024]$
  - Unpolarized intensity:
    $[20, 40, 60, 80, 100, 200, 300, 400, 600, 1024]$
- Building (Fig. 2)
  - AoLP:
    $[60, 120, 180, 240, 300, 360, 420, 480, 600, 1024]$
  - DoLP:
    $[20, 40, 60, 80, 100, 120, 140, 160, 200, 1024]$
  - Unpolarized intensity:
    $[15, 30, 45, 60, 75, 90, 105, 120, 150, 1024]$
- Sunroom (Fig. 3)
  - AoLP:
    $[80, 160, 240, 320, 400, 480, 560, 640, 720, 800, 1024]$
  - DoLP:
    $[15, 30, 45, 60, 75, 90, 105, 120, 150, 1024]$
  - Unpolarized intensity:
    $[15, 30, 45, 60, 75, 90, 105, 120, 150, 1024]$
- Valentines (Fig. 5)
  - AoLP:
    $[60, 120, 180, 240, 300, 360, 420, 480, 600, 1024]$
  - DoLP:
    $[10, 20, 30, 40, 50, 60, 70, 80, 100, 1024]$
  - Unpolarized intensity:
    $[10, 20, 30, 40, 50, 60, 70, 80, 100, 1024]$
- Stream (Fig. 6)
  - AoLP:
    $[120, 240, 360, 480, 600, 700, 800, 900, 950, 1024]$
  - DoLP:
    $[30, 60, 90, 120, 150, 300, 450, 600, 800, 1024]$
  - Unpolarized intensity:
    $[30, 60, 90, 120, 150, 300, 450, 600, 800, 1024]$

For the multiplicative constants $\lambda_{b,i}$ mentioned in Sec. 3.4, we initially set them all to one. We used factors of ten on the fifth and final bands in order to encourage accurate reconstructions for our half and full models.

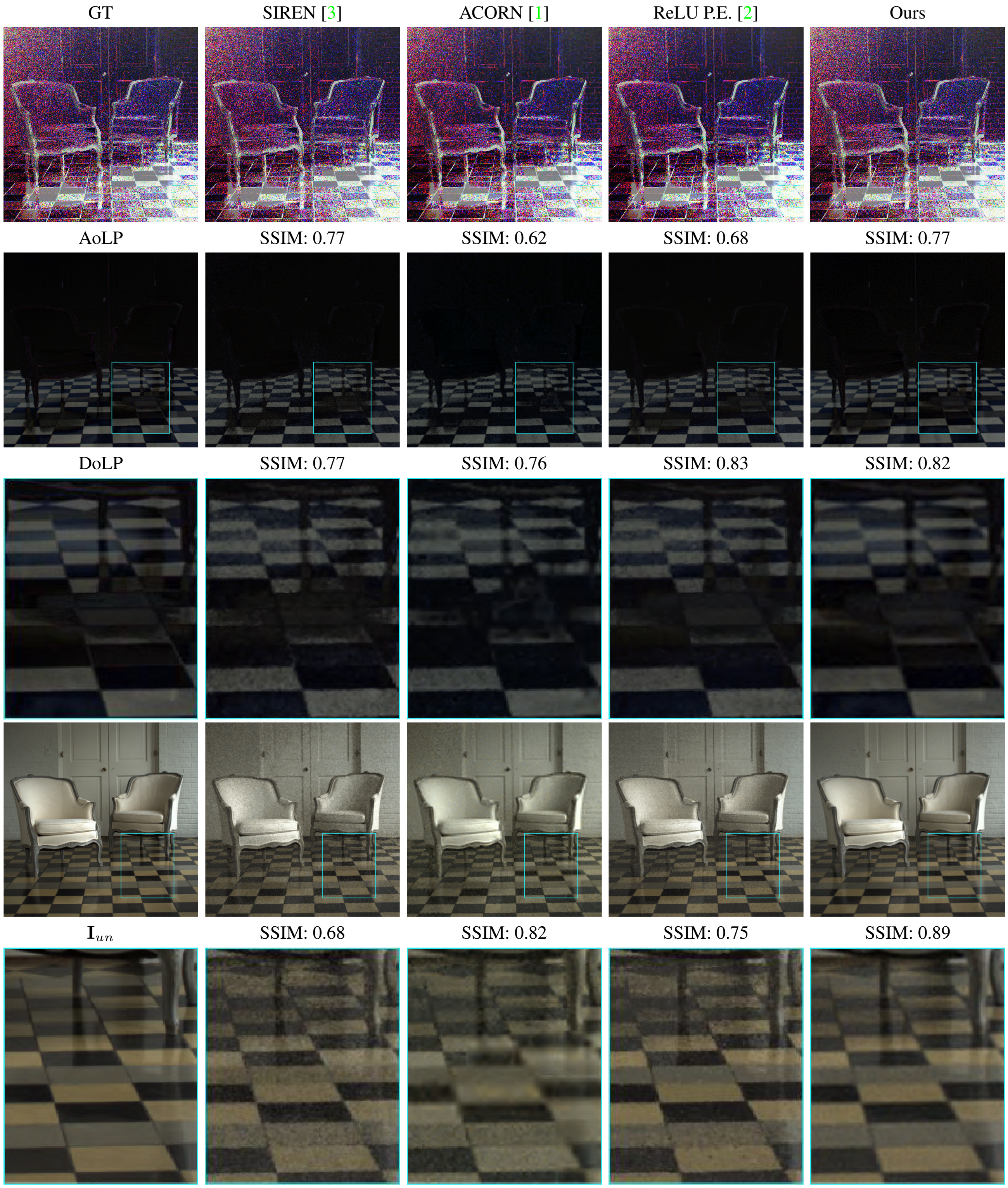

Figure C. **Even with more than double the number of parameters, the baselines still produce lower SSIMs and more artifacts than our model.**

References

[1] 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 Trans. Graph. (SIGGRAPH)*, 40(4), 2021. 2, 3, 4

[2] 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 *European conference on computer vision*, pages 405–421. Springer, 2020. 2, 3, 4

[3] Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. *Advances in Neural Information Processing Systems*, 33:7462–7473, 2020. 2, 3, 4