Supplementary Material: Accurate Polarimetric BRDF for Real Polarization Scene Rendering

Yuhi Kondo, Taishi Ono, Legong Sun, Yasutaka Hirasawa, and Jun Murayama

Sony Corporation, Tokyo, Japan
{Yuhi.Kondo, Taishi.Ono, Legong.Sun, Yasutaka.Hirasawa, Jun.Murayama}@sony.com

1 Experimental Details: Shape from Polarization by CNN

1.1 CNN Architecture

We use the CNN architecture proposed by Zhang et al. [6] with slight modification as follows,

Network Architecture
As shown in Fig. 1, we remove one layer from Zhang et al. citezhang2017physically and replace pooling by strided 2x2 convolution. We also replace concatenation in [6] by simple addition as proposed in [1]. Each weight is initialized by the function of *HeNormal* ([2]). Optimization method is Adam ([3]). Parameters are set as follows: $batchsize = 24$, $weightdecay = 0$ and $learningrate = 0.001$.

Input to CNN
We use intensity, degree of polarization (DoP) and polarization phase for each RGB channel as input. Each input data is scaled to 0–255. For polarization phase, 0°–180° is scaled to 0–255.

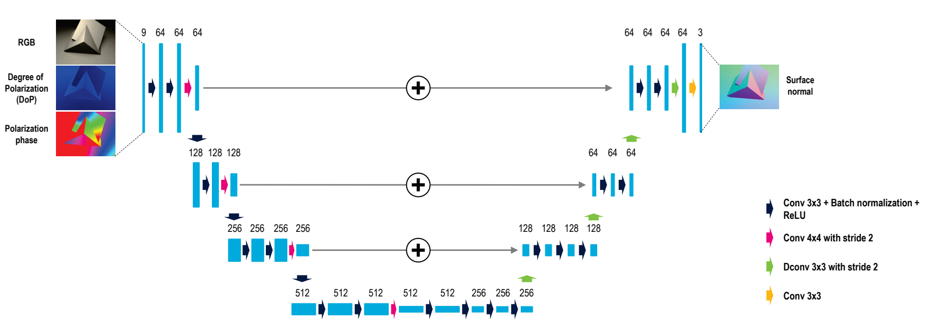

Fig. 1. Proposed CNN architecture: We use U-net [4] as a base architecture. *Dconv* means deconvolution for up-sampling and *ReLU* means Rectified Linear Units.

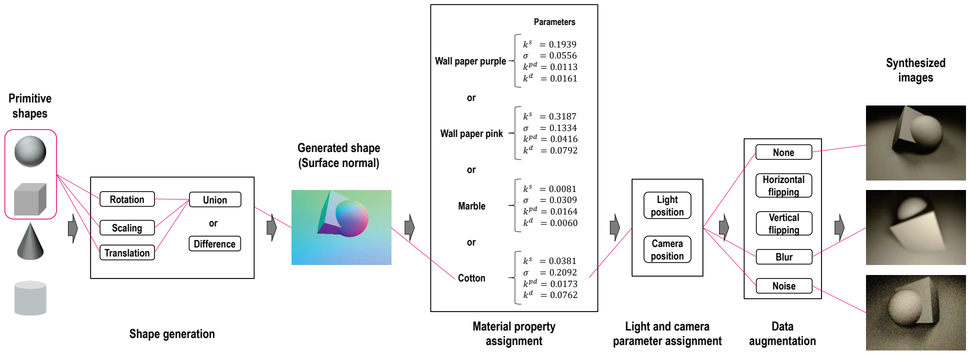

Fig. 2. Dataset generation: The entire process consists of three steps as described in 1.2. Here we show three synthesized images with different augmentation process.

1.2 Dataset

We synthesize 44,305 images with resolution 256 x 192. 35,444 images are used for training and the other 8,861 images are used for validation. The image is composed of the single object and the floor. The entire process of generating dataset is shown in Fig. 2, and each process is described as follows,

Shape Generation: With reference to [5], we generate object shape by combing primitive shapes. There are four types of primitives which are sphere, cube, cone and cylinder. Two of the four primitives are chosen randomly, then they are randomly scaled, rotated and translated. Finally, object shape is generated by taking their union or difference.

Material Property Assignment: We use four materials, two of four have strong specular component and the other have weak specular component. We assign chosen material property randomly to the generated object and floor.

Light and Camera Parameter Assignment: We randomly assign light and camera parameters whose ranges are shown in Table 1.

Data Augmentation: Finally, one of the following process is applied to augment data.

– None
– Horizontal flipping
– Vertical flipping
– Blur: 3x3 averaging filter
– Noise: Gaussian noise with $\sigma = 0.01$

Examples of synthesized images are shown in Fig. 4.

Table 1. Light and camera parameters.

Parameters	Range
light azimuth angle	0°–359°
light zenith angle	30°–70°
number of lights	1–2
camera azimuth angle	0°–359°
camera zenith angle	30°–60°

1.3 Training of CNN

Plot of loss for each epoch is shown in Fig. 3. We have chosen trained parameters of the 30th epoch for evaluation.

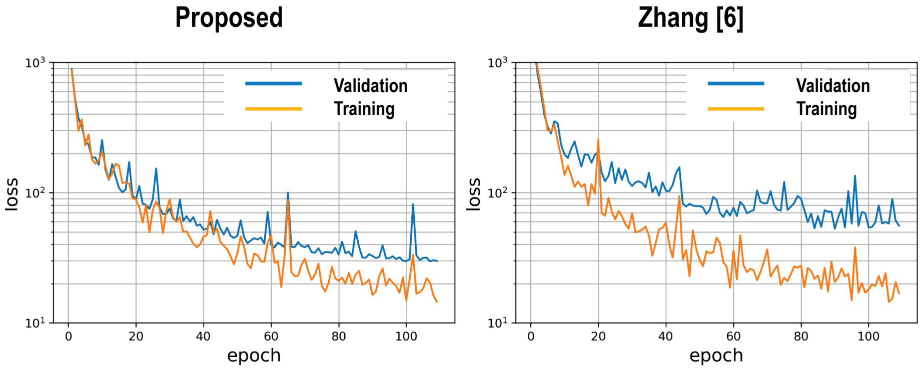

Fig. 3. Plots of loss for the training and validation dataset

Computation time for training is about 50 epochs per day using one GPU (GeForce® GTX 1080 Ti) for both proposed method and Zhang [6].

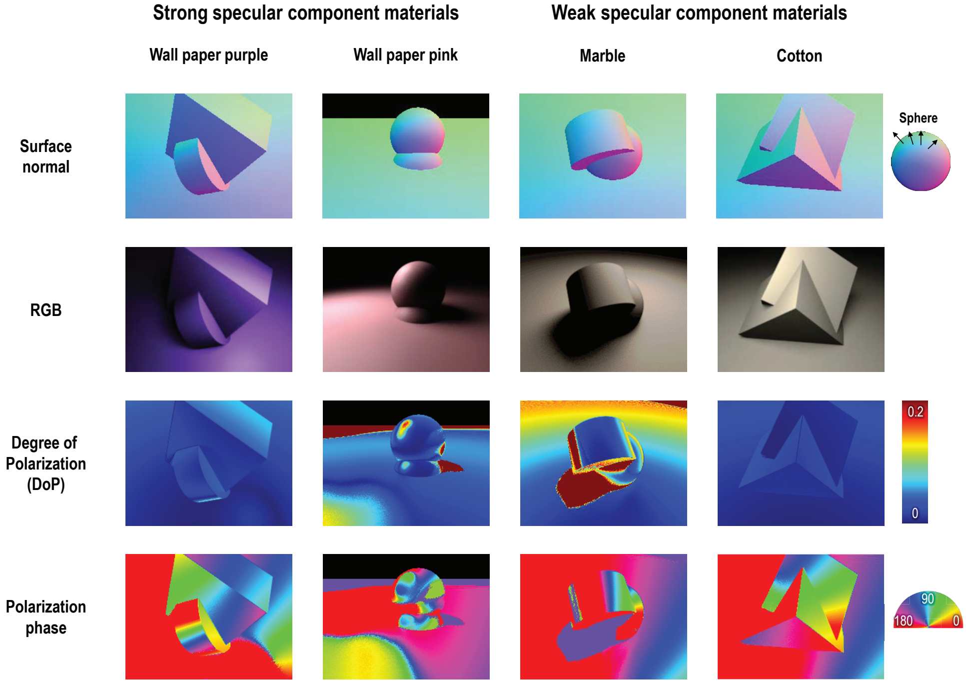

Fig. 4. Example of synthesized data: Surface normal, RGB, DoP, polarization phase images are synthesized. Images in each column are synthesized with chosen shape and material property.

References

1. Ba, Y., Chen, R., Wang, Y., Yan, L., Shi, B., Kadambi, A.: Physics-based neural networks for shape from polarization. arXiv preprint arXiv:1903.10210 (2019)
2. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: IEEE International Conference on Computer Vision. pp. 1026–1034 (2015)
3. Kingma, D., Ba, J.: Adam: A method for stochastic optimization. In: In International Conference on Learning Representations (2015)
4. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)
5. Yang, D., Deng, J.: Shape from shading through shape evolution. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 3781–3790 (2018)
6. Zhang, Y., Song, S., Yumer, E., Savva, M., Lee, J.Y., Jin, H., Funkhouser, T.: Physically-based rendering for indoor scene understanding using convolutional neural networks. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 5287–5295 (2017)