SUPPLEMENTAL MATERIAL: TWO-STEP COLOR-POLARIZATION DEMOSAICKING NETWORK

Vy Nguyen, Masayuki Tanaka, Yusuke Monno, and Masatoshi Okutomi

Tokyo Institute of Technology, Tokyo, Japan

1. SUPPLEMENTAL MATERIAL

This supplemental material presents the experimental results which could not be included in the ICIP main paper due to limited space.

2. EXPERIMENTS

2.1. Datasets and metrics

We evaluated the proposed network with two public available datasets: Tokyo Tech dataset [6,7] and CPDNet dataset [9]. In those two datasets, the ground-truth 12-channel full-color polarization images were taken by the division of time polarimeter approach. Then, the CPFA raw data were synthesized with the corresponding CPFA pattern. We quantitatively evaluated with CPSNR and angle error of AoP following related works [6,7]. For the learning-based methods, each model was trained five times. Then, we evaluate by an average of the metrics.

Tokyo Tech dataset [6,7] includes 40 scenes of 1024 $\times$ 768 resolution of 12-channels full-color polarization images. For each scene, they averaged 1000 data to reduce the effect of noise. The evaluations in this work were conducted on our splits: 30 scenes for the training set, two scenes for the validation set, and eight scenes for the testing set.

CPDNet dataset [9] includes 105 scenes of 1456×1088 resolution, in which 260 scenes are for training, 80 scenes are for validation and 80 scenes are for testing. Yet, there are no public information about how their dataset is splitted. Therefore, we conduct evaluations on our splits according to their published ratio.

2.2. Quantitative comparison

We compared our proposed TCPDNet against other existing methods on Tokyo Tech dataset. We compared with five algorithms; bilinear interpolation, EARI [6], IGRI2 [7], CPDNet [9] (original), and CPDNet [9] (re-trained). The weight of CPDNet (original) is provided by the authors [9], while we re-trained CPDNet (re-trained) with Tokyo Tech dataset and CPDNet dataset (our split). Table 1 and Table 2 show the quantitative comparisons with Tokyo Tech dataset and CPDNet dataset, where the higher CPSNR is better, and the lower angle error is better. We evaluated four RGB color polarization angle images (I_0, I_{45}, I_{90}, and I_{135}), three stokes parameters (S_0, S_1, and S_2), DoP (Degree of Polarization), and AoP (Angle of Polarization). From Table 1 and Table 2, we can find that the proposed TCPDNet clearly outperforms other algorithms.

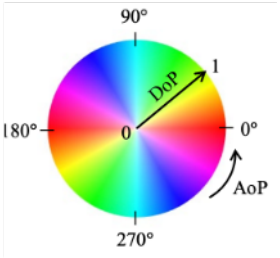

Fig. 1: AoP-DoP visualization color space [6].

2.3. Qualitative comparison

Figure 2 and Figure 3 visualize results of different methods of a scene from Tokyo Tech dataset and CPDNet dataset.

Our proposed TCPDNet is clear from artifacts while existing methods suffer from obvious color artifacts in S0. In the Figure 2's first row example, these methods wrongly estimates the character "S" with red color. In the Figure 2's second row example, CPDNet wrongly estimate red color region as blue. In the Figure 2's third row example, none of these methods can precisely estimate the red line on the top right of the image. In the Figure 3's first row example, EARI and IGRI2 mistakenly estimate the intersection between "N" and the second "A" to be red instead of white, while CPDNet mistakenly estimates brown color in some regions expected to be blue. CPDNet's wrong color estimation is also highlighted in the Figure 3's second and third rows.

Our proposed TCPDNet is also the best in edge retention. In the Figure 2's first row example, the "IS" characters are not properly estimated by EARI and IGRI2. In the Figure 2's second row example, the country name almost disappears in these two method's estimation. In the Figure 3's first and second row example, the text edges estimated by the two method are also blurred. In the Figure 3's third row exaample, the

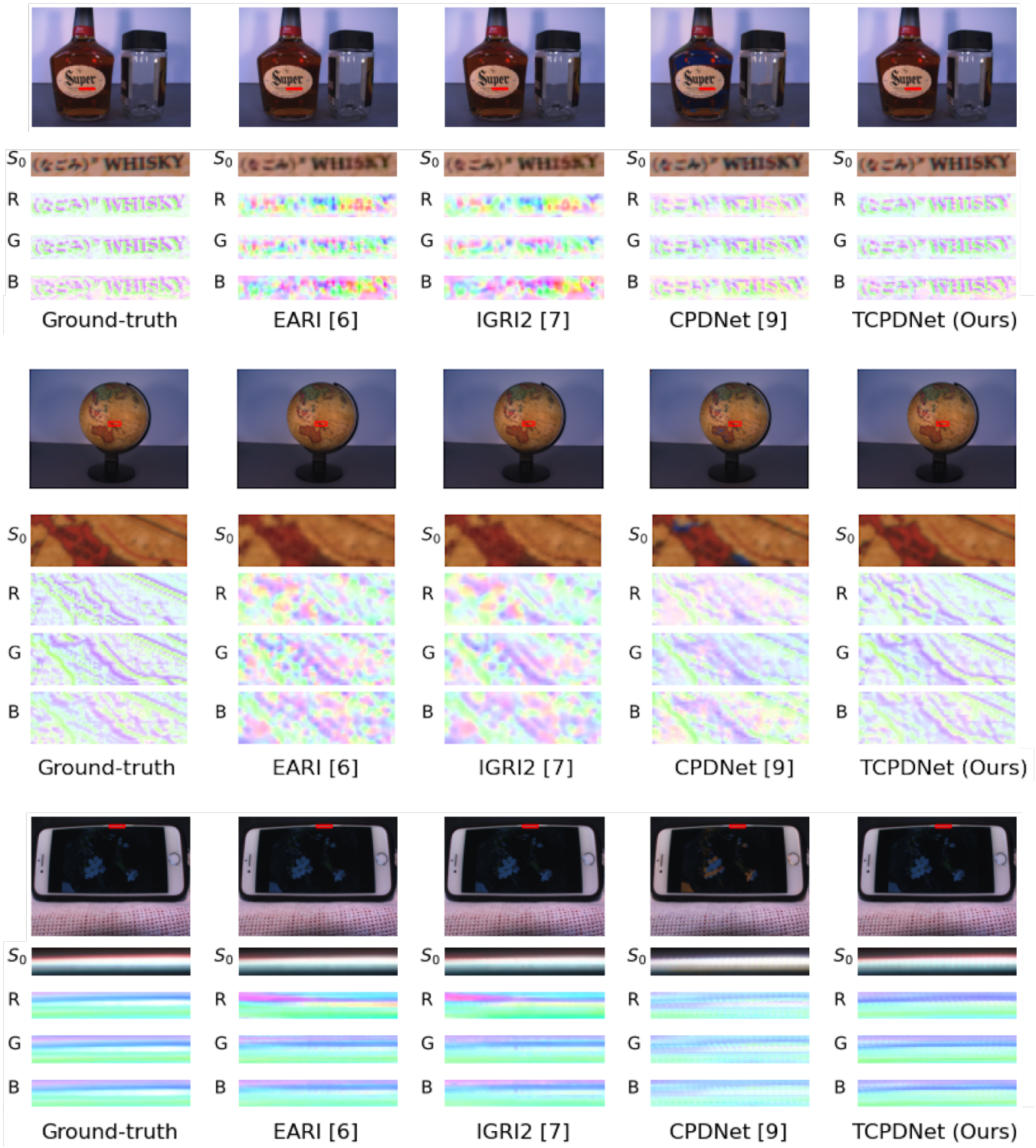

Fig. 2: Qualitative comparison between our proposed TCPDNet and existing methods. The scenes are from Tokyo Tech dataset.

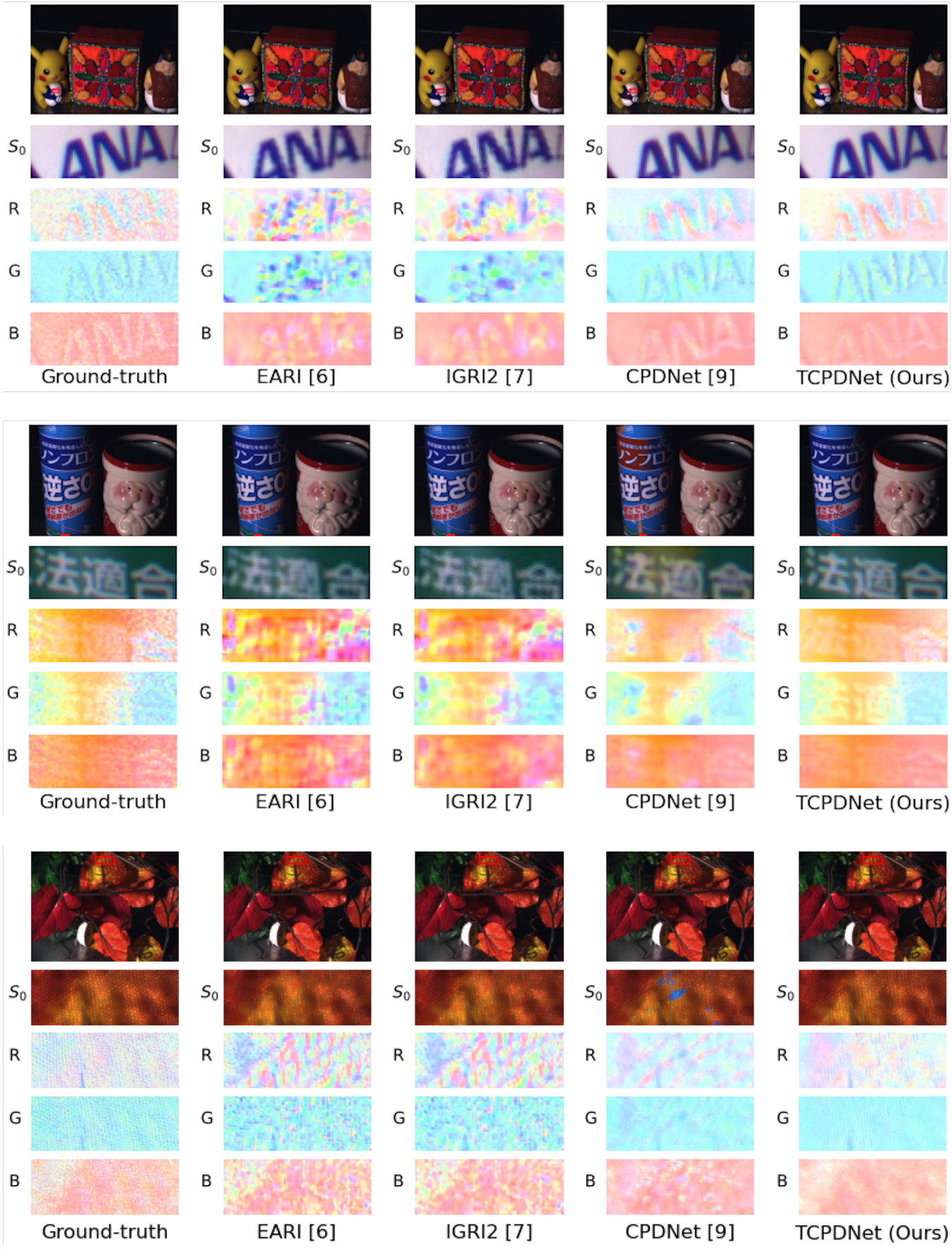

Fig. 3: Qualitative comparison between our proposed TCPDNet and existing methods. The scenes are from CPDNet dataset.

Table 1: Performance comparison on Tokyo Tech dataset.

Method	CPSNR								Angle error
	I_0	I_{45}	I_{90}	I_{135}	S_0	S_1	S_2	DoP	AoP
Bilinear interpolation	34.64	34.27	35.19	34.46	36.01	42.05	39.93	30.33	23.70
EARI [6]	38.33	37.58	39.00	37.77	39.81	45.47	42.82	32.95	20.54
IGRI2 [7]	38.40	37.59	39.07	37.78	39.60	46.38	43.05	33.17	20.05
CPDNet (original) [9]	23.02	24.26	24.33	24.43	24.64	32.35	38.96	24.85	50.42
CPDNet (re-trained) [9]	28.01	27.81	28.10	27.81	28.23	45.23	41.84	31.24	32.32
TCPDNet (Ours)	**43.73**	**43.16**	**44.46**	**43.31**	**44.91**	**52.82**	**48.86**	**38.74**	**12.65**

Table 2: Performance comparison on CPDNet dataset.

Method	CPSNR								Angle error
	I_0	I_{45}	I_{90}	I_{135}	S_0	S_1	S_2	DoP	AoP
Bilinear interpolation	35.11	35.07	34.90	35.03	37.24	40.50	40.62	27.90	29.15
EARI [6]	36.83	36.90	36.79	36.84	39.74	41.51	41.75	28.62	28.91
IGRI2 [7]	36.91	36.97	36.86	36.92	39.41	42.01	42.30	28.99	28.43
CPDNet (pre-trained) [9]	36.92	37.12	36.73	36.80	39.67	41.66	42.01	27.31	31.47
CPDNet (re-trained) [9]	36.70	36.43	36.36	36.19	39.03	41.94	41.99	28.46	30.17
TCPDNet (Ours)	**39.76**	**39.83**	**39.72**	**39.78**	**43.16**	**44.06**	**44.35**	**30.17**	**25.64**

leaf texture estimatinon by the two method looks smoother than expected. CPDNet's estimations are better than hand-crafted methods in general, yet its edge estimations are not as vivid as TCPDNet and suffer from distortion, e.g. the "IS" characters in the Figure 2's first row image.

The differences in AoP-DoP images are drastical amongst different methods. EARI and IGRI2 hardly preserves the edge information. The re-trained CPDNet generally give better estimations but not as close to the ground-truth as TCPDNet.