Simultaneous Acquisition of High Quality RGB Image and Polarization Information using a Sparse Polarization Sensor

Teppei Kurita   Yuhi Kondo   Legong Sun   Yusuke Moriuchi

Sony Group Corporation

{Teppei.Kurita, Yuhi.Kondo, Legong.Sun, Yusuke.Moriuchi}@sony.com
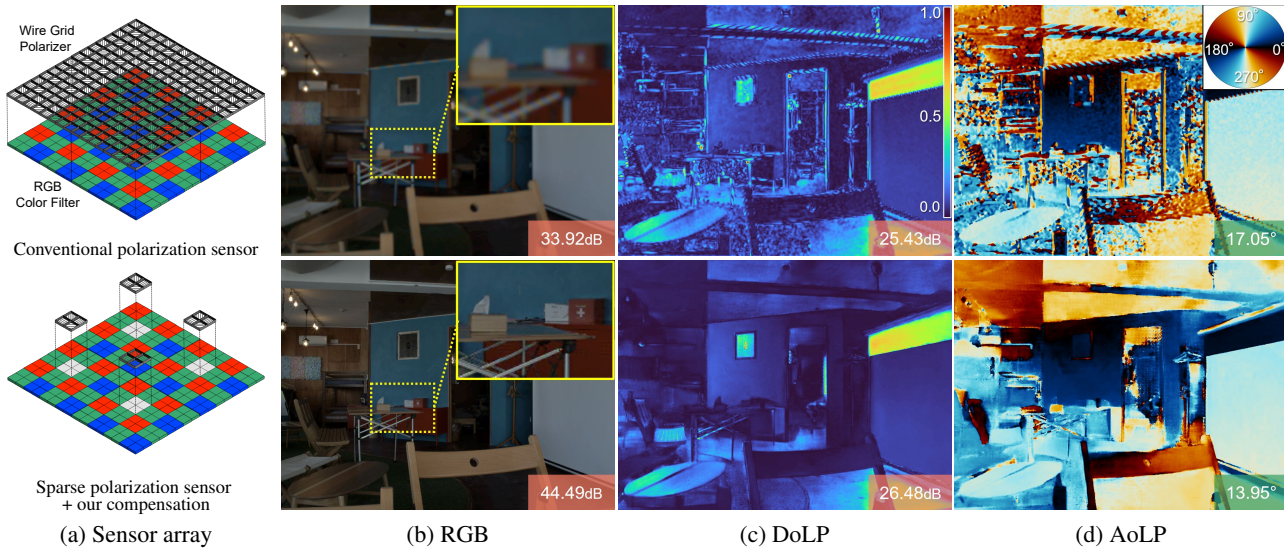
https://github.com/sony/polar-densification

Figure 1. **Compared to conventional polarization sensors, combining a sparse polarization sensor with a compensation method can enable more accurate RGB images and polarization information to be obtained.** The outputs of the conventional polarization sensor (top) and our proposed sensor (bottom). The RGB and degree of linear polarization (DoLP) values indicate the peak signal-to-noise ratio (PSNR) (the larger the better) and the angle of linear polarization (AoLP) values indicate angular error (the smaller the better).

Abstract

This paper proposes a novel polarization sensor structure and network architecture to obtain a high-quality RGB image and polarization information. Conventional polarization sensors can simultaneously acquire RGB images and polarization information, but the polarizers on the sensor degrade the quality of the RGB images. There is a trade-off between the quality of the RGB image and polarization information as fewer polarization pixels reduce the degradation of the RGB image but decrease the resolution of polarization information. Therefore, we propose an approach that resolves the trade-off by sparsely arranging polarization pixels on the sensor and compensating for low-resolution polarization information with higher resolution using the RGB image as a guide. Our proposed network architecture consists of an RGB image refinement network and a polarization information compensation network. We confirmed the superiority of our proposed network in compensating the differential component of polarization intensity

by comparing its performance with state-of-the-art methods for similar tasks: depth completion. Furthermore, we confirmed that our approach could simultaneously acquire higher quality RGB images and polarization information than conventional polarization sensors, resolving the trade-off between the quality of RGB images and polarization information. The baseline code and newly generated real and synthetic large-scale polarization image datasets are available for further research and development.

1. Introduction

The polarization of light represents the orientation of the oscillations of light waves. It can be used to reveal light transport effects [6, 49] such as shape [3, 4], transparency [23, 38, 40], and scattering [54, 60]. In recent years, polarization sensors [56, 16], which can simultaneously acquire the RGB image and polarization information in a single shot by placing the polarizer above the photodi-
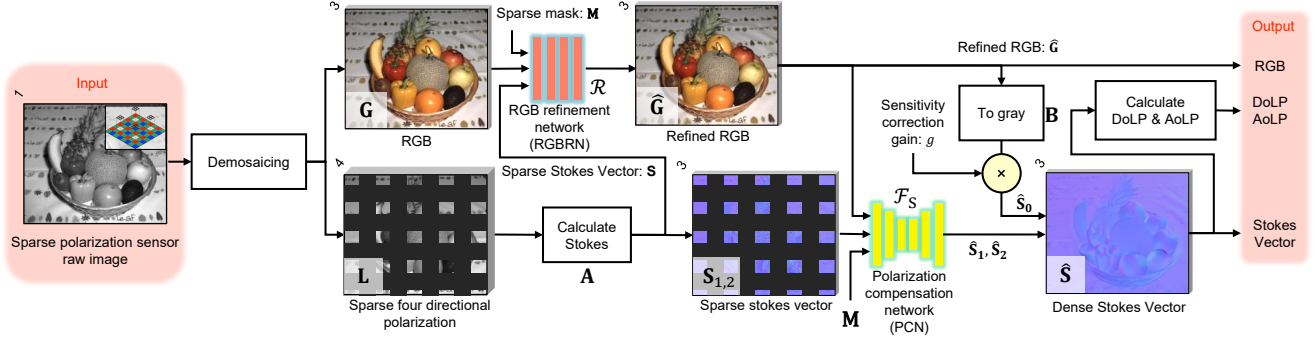
Figure 2. **Proposed stokes network architecture (SNA).** The architecture consists of a refinement network for RGB images and compensation network for polarization information.
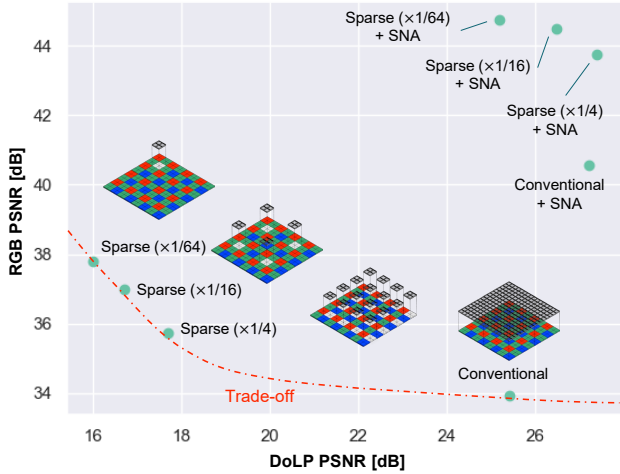

Figure 3. **Plot of PSNR for RGB and DoLP for each method.** Our approach, SNA, resolves the trade-off between RGB image quality and polarization information quality.

ode of the image sensor, as shown at the top in Fig. 1 (a), have become widely used. This widespread implementation has enabled many applications such as shape estimation [15, 41, 22, 44, 7, 63, 61, 14, 20, 64], reflection removal [37, 28, 31, 12], and so on [55, 53, 30, 11, 57, 48]. The acquisition of polarization information is achieved by polarizers placed on the sensor. However, the sensor sensitivity is reduced due to the reduction in light intensity caused by the polarizers. In addition, an unpolarized component is necessary to generate an RGB image, requiring pixel binning (averaging) of polarization components in multiple directions. This reduces the spatial resolution as shown at the top in Fig. 1 (b) (on the other hand, sensitivity is improved by binning). Reducing the degradation of the RGB image by reducing the number of polarization pixels on the sensor, as shown at the bottom in Fig. 1 (a), is feasible in during design and manufacturing, but the resolution of the polarization information is reduced as a side effect.

In this study, we propose an approach to resolve this trade-off between the quality of RGB images and polarization information by sparsely arranging polarization pixels

on the sensor and compensating for low-resolution polarization information with higher resolution using the RGB image as a guide. Our proposed stokes network architecture (SNA) consists of a refining network for RGB images and a compensation network for polarization information using the RGB images of the output of the refining network as a guide, as shown in Fig. 2. The RGB image refinement network helps compensate for polarization information more effectively by correcting demosaicing artifacts and sparse pixels and cleaning up the RGB image used as a guide. The compensation network for polarization information performs compensation only for polarization stokes components S_1 and S_2, which are differential components of polarization intensity. In contrast, the unpolarized component S_0 is generated from the RGB image. As only the difference component is learned, the polarization information can be compensated more stably than the method that compensates for polarization intensity in four directions.

In addition, large amounts of RGB and polarization data are not available to train networks. Although the advent of polarization sensors has reduced the difficulty of acquiring polarization information, significant time and manpower are required to acquire large amounts of data. We used polarization sensors to acquire real-world data and a polarization renderer [26] to generate a large synthetic dataset for training. Our large synthetic dataset is generated using Houdini [1], a 3D software program capable of procedural modeling for the automatic generation of objects. Hence the cost of acquiring a large amount of data is negligible.

We confirmed the superiority of the proposed network architecture by comparing its performance with state-of-the-art (SOTA) methods for similar tasks like depth completion and upsampling. Furthermore, we confirmed that our approach could simultaneously acquire higher quality RGB images and polarization information than conventional polarization sensors, as shown at the bottom in Fig. 1. Specifically, we demonstrated a performance improvement of more than 10 dB and 1 dB in the PSNR for RGB and DoLP images, respectively. Finally, we showed that resolving the trade-off between the quality of RGB images and polariza-

tion information is possible, as shown in Fig. 3. Due to the versatility of our compensation architecture, achieving higher quality output from conventional polarization sensors is also possible. In summary, our contributions are as follows:

- An approach to resolve the trade-off between RGB image quality and polarization information quality by using a sparse polarization sensor and compensation method.

- An effective network architecture for compensating captured polarization information with degradation.

- Providing large-scale polarization datasets that are high-quality real and diverse synthetic.

2. Related work

Several previous studies have proposed denoising [52, 51, 33] or demosaicing [39, 46, 36] to obtain high-quality polarization information. Few research focus on the trade-off between the quality of RGB images and polarization information from polarization sensors. Our strategy of sparsely arranging polarization pixels to avoid the degradation of RGB image quality is derived from image sensors with sparsely arranged phase detection pixels [25]. In these sensors, a portion of the pixels is intentionally light-shielded for image phase detection auto-focus [9]. However, because the RGB image quality is comparable to that of ordinary sensors, these sensors have become popular in recent years in cameras for cell phones and other applications. Therefore, we hypothesized that even if polarization pixels are arranged sparsely, sufficient RGB image quality can be obtained as long as appropriate compensation is provided. We have accordingly conducted research on this hypothesis.

Sparse polarization information has low resolution and needs to be adequately compensated using the RGB image as a guide. Similar problems are encountered during the super-resolution of hyperspectral images, and depth completion and upsampling. Because hyperspectral images often have low spatial resolution due to the adverse effects of dense spectral sampling, a lot of research on fusing hyperspectral images with high-resolution RGB images to produce high-resolution hyperspectral images has been conducted. However, many methods [27, 13] use constraints related to the unique physical properties of hyperspectral images and are not versatile enough to be applied directly to other problems. In addition, depth completion [19, 10, 58, 62, 35, 17, 45] or upsampling [24, 34] is the task of taking a sparse depth map or low-resolution depth image and transforming it into high-resolution by referencing to an RGB image. This method has been the subject of much research, particularly using neural networks. As the demand for autonomous driving, augmented reality,

(a) Pixel binning   (b) Regular pixel   (c) Polarization pixel
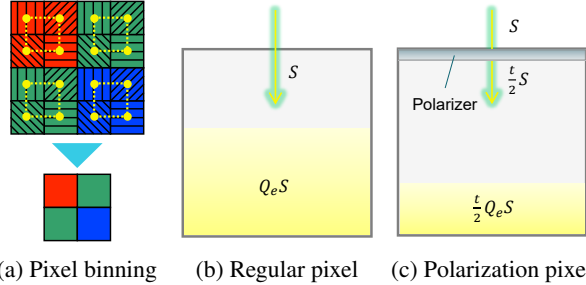
Figure 4. **Pixel binning and sensitivity.** (a) Pixel binning (averaging) in conventional polarization sensors. Four pixels of the same color in a neighborhood are averaged to generate the unpolarized component, reducing resolution but increases sensitivity. (b,c) Polarization pixel is less sensitive because the polarizer reduces the amount of light.

gesture recognition, etc., has increased, these fields have become more competitive, and SOTA networks [18, 43, 50] have become more sophisticated. However, because these methods are specialized for depth estimation, they produce many artifacts when applied directly for polarization information compensation. Hence, in this study, we propose a new network architecture suitable for compensating polarization information.

3. Method

This section details the structure of the proposed sensor, problem formulation, network architecture for obtaining high-quality RGB images and polarization information, and newly acquired and generated real-world and synthetic datasets.

3.1. Sparse polarization sensor

An example of a sparse polarization sensor is shown at the bottom in Fig. 1 (a). This is a structure where four polarization pixels of different angles are arranged in an 8×8 area in a Quad Bayer array RGB sensor. The proportion of polarized pixels is r, where $r = 1/16$. A white color filter is placed in the polarization pixel area to increase the sensitivity of the sensor to the visible light wavelength band.

Spatial resolution: To produce an RGB image using a conventional polarization sensor, pixel binning (averaging) of the four polarization angles must be done to create an unpolarized component, as shown in Fig. 4 (a). Hence, if the total number of pixels is N, the number of RGB image pixels in a conventional polarization sensor would be $N/4$. In contrast, the sparse polarization sensor can output regular pixels without binning, and the number of pixels would be $(1 - r)N$. Thus, the sparse polarization sensor has $4(1 - r)$ times more pixels in the RGB image than the conventional polarization sensor, translating to 3.75 times more pixels when $r = 1/16$. As the sparse polarization sensor has fewer
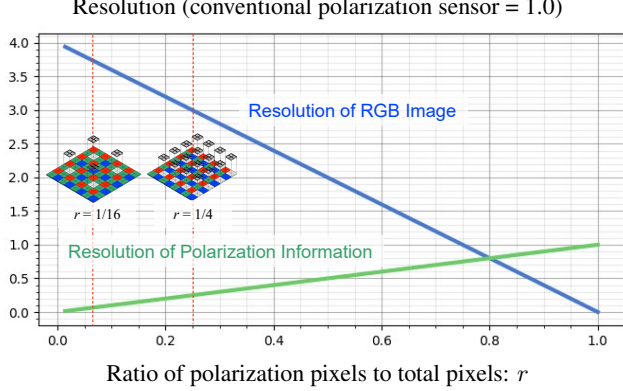
Figure 5. **Resolution of sparse polarization sensors relative to conventional polarization sensors.** There is a trade-off between the resolution of the RGB image and the resolution of polarization information.

polarization pixels, the number of pixels for polarization information is r times that of a conventional polarization sensor. Hence, as shown in Fig. 5, the relationship between the resolution of the RGB image and the resolution of polarization information for the percentage of polarized pixels r is a trade-off.

Sensitivity (signal-to-noise ratio (SNR)): Conventional polarization sensors generate RGB images from polarized pixels, whereas sparse polarization sensors generate RGB images from regular pixels. As shown in Fig. 4 (b,c), when unpolarized light enters a polarized pixel, the light intensity becomes $t/2$ times that of a regular pixel, where t is the transmittance of the polarizer ($0 \leq t \leq 1$). For simplicity, assume that the noise factor, read noise and transmittance of the color filter are negligible, and the number of photons and quantum efficiency are S and Q_e, respectively. Accordingly, the SNR of a regular pixel and polarized pixel is $\sqrt{Q_e S}$ and $\sqrt{t Q_e S/2}$, respectively. When generating RGB images, the average SNR for the conventional polarization sensor is $\sqrt{2 t Q_e S}$ because it averages four polarization pixels, while the average SNR for the sparse polarization sensor is $\sqrt{Q_e S}$. Therefore, the sparse polarization sensor has $\sqrt{1/2t}$ times higher SNR for the RGB image than the conventional polarization sensor. When the transmittance is set to $t = 0.7$, the standard value for polarization sensors, the SNR is slightly lower at approximately 0.85 times. Conversely, the sparse polarization sensors is slightly more sensitive to polarization information due to the white color filter.

Table 1 summarizes the characteristics of sparse polarization sensors compared to conventional polarization sensors. The more sparse the polarization information, the better the resolution of the RGB image, whereas the resolution of the polarization information deteriorates. Conversely, the sensitivity of the RGB image is slightly reduced and the sensitivity of the polarization information is slightly increased.

Table 1. **Characteristics of sparse polarization sensors compared to conventional polarization sensors.** r is the ratio of the number of polarization pixels to the total number of pixels ($0 \leq r \leq 1$) and t is the transmittance of the polarizer ($0 \leq t \leq 1$).

	RGB	Polarization
Resolution	$\times 4(1-r)$	$\times r$
Sensitivity (SNR)	$\times \sqrt{1/2t}$	Slightly better

Regardless, these changes are trivial compared to the spatial resolution, on which the quality of the final RGB image and polarization information highly depend.

3.2. Problem formulation

The aim of this research is to generate a compensated stokes component $\hat{\mathbf{S}} = [\hat{\mathbf{S}}_0, \hat{\mathbf{S}}_1, \hat{\mathbf{S}}_2] \in \mathbb{R}^{M \times N \times 3}$ for a high-resolution RGB image $\mathbf{G} \in \mathbb{R}^{M \times N \times 3}$ and sparse four-polarization image $\mathbf{L} = [\mathbf{L}_0, \mathbf{L}_{45}, \mathbf{L}_{90}, \mathbf{L}_{135}] \in \mathbb{R}^{M \times N \times 4}$ filled with zeros except for the polarization pixels. M and N are the height and width of the image, respectively. As the polarization sensor cannot acquire the circular polarization component, $\mathbf{S}_3$ is omitted. The transformation matrix $\mathbf{A}$ transforms each component of the four polarization angles into its corresponding stokes vector as follows:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_0 \\ \mathbf{S}_1 \\ \mathbf{S}_2 \end{bmatrix} = \mathbf{A} \cdot \mathbf{L} = \begin{bmatrix} (\mathbf{L}_0 + \mathbf{L}_{45} + \mathbf{L}_{90} + \mathbf{L}_{135})/4 \\ (\mathbf{L}_0 - \mathbf{L}_{90})/2 \\ (\mathbf{L}_{45} - \mathbf{L}_{135})/2 \end{bmatrix} \quad (1)$$

A naive solution is to learn a mapping $\mathcal{F}_L : \mathbf{L} \to \hat{\mathbf{L}}$ such that $\mathbf{G}$ is used as a guide to produce a high-quality four-polarization angle image $\hat{\mathbf{L}}$, and then convert $\hat{\mathbf{L}}$ to a stokes vector $\hat{\mathbf{S}}$. This conversion is carried out as follows:

$$\hat{\mathbf{S}} = \mathbf{A} \cdot \mathcal{F}_L(\mathbf{L}, \mathbf{G}; \theta_{\mathcal{F}_L}), \quad (2)$$

where $\theta_{\mathcal{F}_L}$ is the set of learned weights. However, this method does not preserve the minute relationship between the four polarization angles well and produces significant artifacts.

Another basic approach is to learn the mapping $\mathcal{F}_S : \mathbf{S} \to \hat{\mathbf{S}}$ such that instead of compensating for the four polarization angles directly, $\mathbf{L}$ is transformed into a sparse stokes vector $\mathbf{S}$ to produce a high-quality stokes vector $\hat{\mathbf{S}}$, as follows:

$$\hat{\mathbf{S}} = \mathcal{F}_S(\mathbf{A} \cdot \mathbf{L}, \mathbf{G}; \theta_{\mathcal{F}_S}). \quad (3)$$

In this approach, components such as $\mathbf{S}_1$ and $\mathbf{S}_2$, primarily low-frequency and minor values, are recovered with relatively good quality. Regardless, the $\mathbf{S}_0$ component, with many high-frequency textures, is significantly degraded. Another problem is that the artifacts primarily caused by sparse polarization pixels in the RGB image $\mathbf{G}$ degrades compensation performance, particularly near the edges.

Therefore, we propose the generation of the $\mathbf{S}_0$ component from the RGB image and the mapping $\mathcal{F}_S$ to be performed using only the difference (polarization) components of the stokes vector, $\mathbf{S}_1$ and $\mathbf{S}_2$. The $\mathbf{S}_0$ component of the stokes vector represents the unpolarized component. Therefore, as long as the sensitivity difference between pixels is absorbed, a higher quality image can be generated from the higher resolution unpolarized RGB image without generating from polarized pixels. We also correct the demosaicing artifacts in the RGB image to produce a higher quality RGB image $\hat{\mathbf{G}}$ to be used as a guide for polarization information compensation. Such a process is formulated as follows:

$$\mathbf{S} = \mathbf{A} \cdot \mathbf{L}, \; \hat{\mathbf{G}} = \mathcal{R}(\mathbf{G}, \mathbf{S}, \mathbf{M}; \theta_{\mathcal{R}}),$$
$$\hat{\mathbf{S}}_0 = g \cdot \mathbf{B} \cdot \hat{\mathbf{G}}, \quad (4)$$
$$\hat{\mathbf{S}}_{1,2} = \mathcal{F}_S(\mathbf{S}_1, \mathbf{S}_2, \hat{\mathbf{G}}, \mathbf{M}; \theta_{\mathcal{F}_S}),$$

where the mapping $\mathcal{R} : \mathbf{G} \to \hat{\mathbf{G}}$ compensates for RGB, g is a gain to absorb the sensitivity difference between the unpolarized and polarized pixels, $\mathbf{B}$ is a transformation matrix that converts the pixels from RGB to grayscale, and $\mathbf{M}$ is a mask for polarized pixels (a binary image with 1 for pixels with and 0 for pixels without polarization).

3.3. Network architecture

This study proposes an end-to-end network architecture, SNA, that compensates for the low-resolution polarization pixels using RGB images, as shown in Fig. 2. First, taking the RAW data from the sparse polarization sensor as input, a demosaicing process is performed to separate data into a three-channel RGB image and a four-channel image with four sparse polarization angles. The sparse four-polarization angle image is converted to a stokes vector by the transformation matrix $\mathbf{A}$. Next, the RGB image refinement network corrects the demosaicing artifacts in the RGB image. Then, using the refined RGB image and polarization components of the sparse stokes components $\mathbf{S}_{1,2}$ as inputs, the polarization components of the high-resolution stokes components are estimated by a compensation network of polarization information. After the refined RGB image is grayscaled and gain applied to absorb sensitivity differences, high-resolution $\mathbf{S}_0$ component is calculated to obtain the final stokes vector and RGB image. The proposed SNA provides a compensation of $\mathbf{S}_{1,2}$, which is one of the main contributions of this study. Compared to the basic compensation architecture that uses four polarization intensity images that are not separated into $\mathbf{S}_{1,2}$, SNA can compensate efficiently under polarization constraints (Tab. 4).

RGB refinement network (RGBRN): Sparse polarization sensors have polarization pixels interspersed with regular RGB sensors. Therefore, demosaicing may cause artifacts in that polarized pixel region. Such demosaicing artifacts in RGB images can interfere with the compensation of
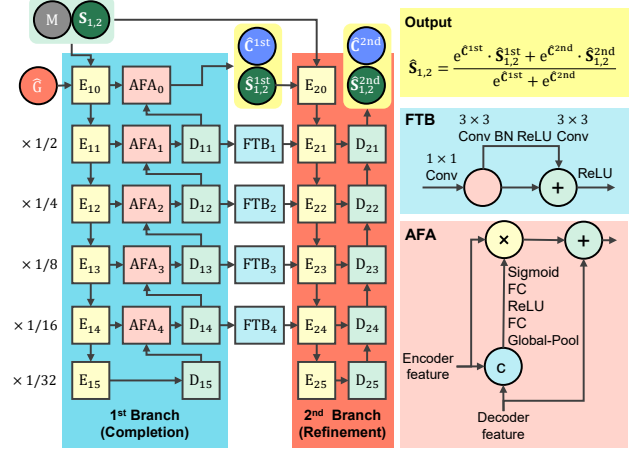

Figure 6. **Polarization compensation network (PCN).** The network consists of one branch that complements sparse polarization information using the RGB image as a guide and a second branch that refines the obtained dense polarization information.

polarization information. Hence, we use a refinement network to clean up the RGB image. The network architecture is based on [21], with a modification in residual learning to stabilize learning. In addition to the RGB image, sparse polarization components $\mathbf{S}_{0,1,2}$ and a mask $\mathbf{M}$ are used as input to complement the information in the missing polarization pixels, thereby improving refinement.

Polarization compensation network (PCN): As previously mentioned, we compensate only for polarization information $\mathbf{S}_{1,2}$. $\mathbf{S}_{1,2}$, the refined RGB image $\hat{\mathbf{G}}$, and mask $\mathbf{M}$ are taken as input, and compensation is performed by a two-branch network shown in Fig. 6, based on the ENet, the backbone in [19]. The network first generates dense polarization information $\hat{\mathbf{S}}_{1,2}^{1st}$, and confidence levels $\mathbf{C}^{1st}$ as the first output, using the RGB image as a guide in the first branch. The branch is intended to complement the polarization information. Next, the second blanch is used to refine the polarization information. The dense polarization information $\hat{\mathbf{S}}_{1,2}^{2nd}$ and confidence level $\mathbf{C}^{2nd}$ are generated as the second output using the first output $\hat{\mathbf{S}}_{1,2}^{1st}$ and sparse $\mathbf{S}_{1,2}$ as inputs. Finally, the first and second outputs are blended at their respective confidence levels $\hat{\mathbf{C}}$ to obtain polarization information $\hat{\mathbf{S}}_{1,2}$ as the final output.

Because ENet is intended to complement depth, it cannot be applied to complement polarization information. We expand the number of input and output channels to two and extract the ReLU just before the output to allow negative results. The ENet has a direct skip connection of features from the first branch decoder to the same resolution layer of the second branch encoder. However, features suitable for second branch refinement are not always generated by the first branch for completion purposes. This gap causes performance degradation. Therefore, we use a feature transfer

Table 2. Comparison among different polarization datasets

Dataset	Collection	Size	Resolution
Ba [5]	Polarization camera	263	1.3M
Lei [29]	Polarization camera	522	1.3M
Ono [42]	Polarization camera	69	5.0M
	Polarizer rotation	13	4.3M
	Synthetic	11000	1.3M & 0.4M
Ours	Polarization camera	811	5.0M
	Polarizer rotation	238	20.0M

(a) Auto-generated floor and objects   (b) Example of data

Figure 7. **Automatically generated synthetic dataset.** (a) Floors and objects are generated procedurally. (b) Cameras are positioned randomly to acquire a wide variety of data at no cost.

block (FTB) [59, 32] to transfer the features generated by the first branch into a form suitable for the second branch. In addition, the first branch of the ENet extracts RGB features through an encoder. It compensates $S_{1,2}$ with the decoder, so different types of features are directly added with skip connections, causing a limitation in the expressive power of the network. We apply an attention-based feature aggregation block (AFA) [59, 32] that takes into account the global and visual features of the scene, assigns higher weights to important channels on the encoder side, and adds them to the decoder side features. This enables learning of more flexible representations in the first branch.

Loss function: We use L1 loss for $S_{1,2}$ and L2 loss for RGB G to train our network. In the early stages of learning, the intermediate outputs, $S_{1,2}^{1st}$ and $S_{1,2}^{2nd}$, are also supervised, defined by

$$\mathcal{L}_S(\hat{S}) = ||\hat{S}_{1,2} - S_{1,2}^{gt}||_1, \ \mathcal{L}_G(\hat{G}) = ||\hat{G} - G^{gt}||_2$$
$$\mathcal{L} = \mathcal{L}_S(\hat{S}) + \lambda\{\mathcal{L}_S(\hat{S}^{1st}) + \mathcal{L}_S(\hat{S}^{2nd})\} + \mathcal{L}_G(\hat{G}) \quad (5)$$

where $S_{1,2}^{gt}$, G^{gt} are the ground truth, and λ is the hyperparameter, a weight that decreases with the number of epochs.

3.4. Dataset

We obtained real-world RGB and polarization information datasets using two methods. One method uses a polarization sensor, which is less challenging to acquire data, but the data quality is inferior. The other method is to rotate the polarizer in front of a regular RGB camera to acquire the data, which has higher quality but costs much more. Hence, we generated a large synthetic dataset with our polarization renderer to compensate for these shortcomings. Our approach uses Houdini to procedurally generate 3D objects and indoor scenes, as shown in Fig. 7. Accordingly large amounts of data can be obtained at no cost. We used rule-based floorplan generation [8], with object cameras randomly placed within the rules, and object textures and bump maps obtained from the Unreal Engine marketplace [2]. Table 2 shows a comparison of our real-world and synthetic datasets with existing datasets. Our datasets outperformed both real-world and synthetic in their data size, and we acquired many high-resolution datasets.

4. Experiments

4.1. Dataset and implementation details

To evaluate our method, we performed numerous experiments on our dataset. We used 10729 images (10000 synthetic and 729 real-world images acquired with an FLIR BFS-U3-51S5P-C color polarization sensor) for training and 1082 images (1000 synthetic and 82 real-world images) for validation. For evaluation, we used 238 real-world images acquired by rotating the polarizer in front of the RGB camera FLIR BFS-U3-200S6C-C. Raw data from a sparse polarization sensor with a resolution of 768×576 was generated by performing simulations on the dataset. The transmittance of the polarizer was set to $t = 0.7$. We added shot noise with a standard deviation of $F_n\sqrt{S}$ to the real-world evaluation data based on a noise model taken from a real sensor, where S is the pixel value and F_n is the noise factor. We used $F_n = 0.72$ in our experiments unless otherwise stated, as it is the value for 0 dB analog gain on the real sensor.

Our network architecture was implemented in PyTorch on a PC equipped with an NVIDIA A100 GPU. We performed 30 epochs of network training with a batch size of 5, an initial learning rate of 0.001 (gradual decay), and adopted the model of the epoch with the best validation results. Using the loss defined in Eqn. 5, $\lambda = 0.2$ was set for the initial epoch and gradually reduced to 0.

4.2. Assessment

We use the PSNR and structural similarity index (SSIM) for the RGB images, root-mean-square error (RMSE) for the entire stokes vector $S_{0,1,2}$ and polarization component, only $S_{1,2}$ for polarization information, PSNR of the DoLP $\sqrt{s_1^2 + s_2^2}/s_0$, and angular difference of the AoLP $\frac{1}{2}\tan^{-1}(s_2/s_1)$, s represents the pixel value of S. We trained the model five times with different random seeds and used the average of the evaluation results. The processing speed of our method was 0.05 seconds per 768×576 image. Due to space limitations in the paper, a detailed evaluation
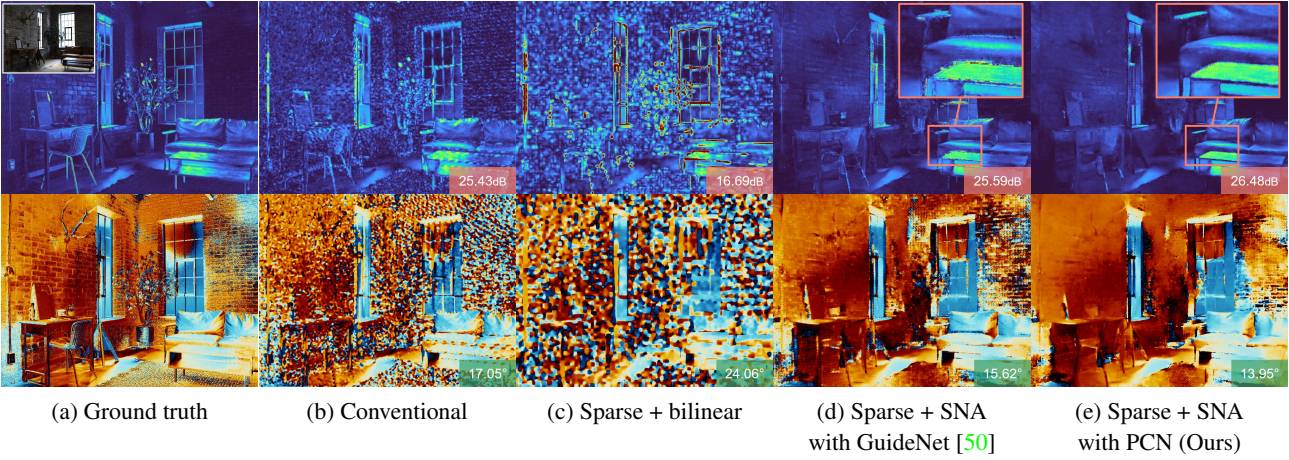
Figure 8. **Comparison with other methods.** Evaluation at $r = 1/16$. The top is DoLP and the bottom is AoLP. (d) shows the results obtained after replacing our PCN with GuideNet in the proposed stokes network architecture (SNA) described in Sec. 3.3.

Table 3. **Comparison with the basic methods (Sec. 3.2).** r is the percentage of polarized pixels.

r	Method	$\mathbf{S}_{0,1,2}$ RMSE ↓ [$\times 10^{-3}$]	$\mathbf{S}_{1,2}$ RMSE ↓ [$\times 10^{-3}$]	DoLP PSNR ↑ [dB]	AoLP Error ↓ [°]
$\frac{1}{4}$	Eqn. 2	11.016	10.632	16.73	31.21
	Eqn. 3	20.622	4.059	26.69	12.54
	Ours	**4.881**	**3.824**	**27.41**	**12.36**
$\frac{1}{16}$	Eqn. 2	9.103	8.955	19.00	27.49
	Eqn. 3	27.562	4.353	25.40	**13.57**
	Ours	**4.707**	**4.151**	**26.48**	13.95
$\frac{1}{64}$	Eqn. 2	8.743	8.759	17.56	32.88
	Eqn. 3	26.833	4.996	23.88	15.74
	Ours	**5.032**	**4.801**	**25.19**	**15.56**

Table 4. **Ablation study (Sec. 3.3).** Evaluation at $r = 1/4$.

Operation	$\mathbf{S}_{0,1,2}$ RMSE ↓ [$\times 10^{-3}$]	$\mathbf{S}_{1,2}$ RMSE ↓ [$\times 10^{-3}$]	DoLP PSNR ↑ [dB]	AoLP Error ↓ [°]
Baseline	10.948	10.574	16.57	31.70
+ SNA	10.328	4.046	26.94	**12.31**
+ RGBRN	4.974	3.979	26.75	13.43
+ FTB	4.952	3.952	26.81	13.14
+ AFA (Ours)	**4.881**	**3.824**	**27.41**	12.36

has been provided in the supplementary material.

Comparison with the basic methods (Sec. 3.2): We first confirmed the effectiveness of our method (Eqn. 4) by comparing it with Eqn. 2 and Eqn. 3. Table 3 shows that for three different r (1/4, 1/16, and 1/64), our method provides the most accurate results for polarization information. Eqn. 3 has some good results, but the RMSE of $\mathbf{S}_{0,1,2}$ is extremely poor, indicating that $\mathbf{S}_0$ is not well estimated.

Ablation study: Table 4 shows the ablation study confirming the validity of our network architecture described in

Table 5. **Comparison of the results obtained after replacing PCN with depth completion and upsampling networks in our network architecture (Sec. 3.3).** Evaluation at $r = 1/16$.

Method	$\mathbf{S}_{0,1,2}$ RMSE ↓ [$\times 10^{-3}$]	$\mathbf{S}_{1,2}$ RMSE ↓ [$\times 10^{-3}$]	DoLP PSNR ↑ [dB]	AoLP Error ↓ [°]
UNet [47]	4.974	4.568	25.00	16.31
U2Net [44]	5.537	5.224	23.90	19.11
FDSR [18]	5.128	4.837	25.12	15.48
NLSPN [43]	4.905	4.470	23.97	20.20
GuideNet [50]	4.859	4.390	25.59	15.62
PCN (Ours)	**4.707**	**4.151**	**26.48**	**13.95**

Table 6. **Comparison of real-world (R) and synthetic (S) datasets (Sec. 3.4).** Evaluation at $r = 1/4$.

Data	Train size	$\mathbf{S}_{0,1,2}$ RMSE ↓ [$\times 10^{-3}$]	$\mathbf{S}_{1,2}$ RMSE ↓ [$\times 10^{-3}$]	DoLP PSNR ↑ [dB]	AoLP Error ↓ [°]
R	729	11.162	5.240	25.40	14.79
S	729	6.346	5.342	24.05	15.55
S	10000	5.137	4.093	26.58	13.11
R+S	10729	**4.881**	**3.824**	**27.41**	**12.36**

Sec. 3.3. Baseline is a method that compensates for four polarization intensity images. The effectiveness of each measure—particularly, the polarization-constrained $\mathbf{S}_{1,2}$ compensation (SNA) and RGBRN—is verified.

Comparison with depth completion networks: We also applied a general-purpose network (UNet [47], U2Net [44]) and SOTA networks for depth completion and upsampling, GuideNet [50], NLSPN [43], and FDSR [18], and compared their performance with our PCN in Sec. 3.3. Each network was modified for polarization, and all conditions were identical except for PCN. In other words, among the architectures proposed in Sec. 3.3, $\mathbf{S}_{1,2}$ compensation and RGBRN were adopted and the difference between our PCN and other networks was evaluated. Table 5 shows the quan-

Table 7. **Comprehensive evaluation of conventional and sparse polarization sensors, taking noise into account.** F_n is the noise factor (higher is noisier).

F_n	Polarization sensor	r	Method	$\mathbf{S}_{0,1,2}$ RMSE ↓ [$\times 10^{-3}$]	$\mathbf{S}_{1,2}$ RMSE ↓ [$\times 10^{-3}$]	DoLP PSNR ↑ [dB]	AoLP Error ↓ [°]	RGB PSNR ↑ [dB]	RGB SSIM ↑
0.72	Conventional	0	-	15.124	6.629	25.43	17.05	33.92	0.9251
			Ours	6.915	3.854	27.22	**12.35**	40.56	0.9749
	Sparse	$\frac{1}{4}$	Bilinear	14.643	13.438	17.68	23.34	35.74	0.9579
			Ours	4.881	**3.825**	**27.41**	12.36	43.76	0.9878
		$\frac{1}{16}$	Bilinear	13.518	13.579	16.69	24.06	36.98	0.9690
			Ours	**4.707**	4.151	26.48	13.95	44.49	0.9897
		$\frac{1}{64}$	Bilinear	13.369	13.973	15.98	25.10	37.79	0.9732
			Ours	5.032	4.802	25.19	15.56	**44.74**	**0.9902**
3.6	Conventional	0	-	15.216	6.876	23.77	23.99	33.85	0.9233
			Ours	7.238	**3.972**	**26.06**	**14.93**	39.79	0.9685
	Sparse	$\frac{1}{4}$	Bilinear	15.131	13.977	15.90	30.90	35.38	0.9493
			Ours	5.617	4.067	25.38	17.05	41.83	0.9759
		$\frac{1}{16}$	Bilinear	14.064	14.100	15.18	31.37	36.47	0.9592
			Ours	**5.380**	4.356	25.22	18.23	42.49	0.9789
		$\frac{1}{64}$	Bilinear	13.954	14.522	14.57	32.05	37.18	0.9630
			Ours	5.690	5.040	23.86	22.51	42.73	0.9798

titative comparison results and Fig. 8 shows the qualitative comparison of GuideNet. Each result shows that our PCN can compensate for polarization information better than the other networks.

Dataset (Sec. 3.4): We also evaluated the training performance of each of our datasets (real-world and synthetic), confirming the effectiveness of the large synthetic dataset, as shown in Tab. 6. The model can learn just as well from the synthetic dataset as the real-world dataset (with significant improvement in accuracy for $\mathbf{S}_{0,1,2}$) for the same number of training images. A higher number of training images can produce even higher quality results. In addition, a mixture of real-world and synthetic datasets performs best because the domain gap is eliminated.

Comprehensive noise-aware evaluation: To further investigate noise immunity, we reproduced situations of high noise factor F_n (high analog gain) from the noise model of the actual sensor and performed comprehensive evaluation of the performances of the conventional and sparse polarization sensors at various r, as shown in Tab. 7. Observing the results of the evaluation, we see that our compensation method remains effective even in noisy situations ($F_n = 3.6$). A visual representation of the evaluation results at $F_n = 0.72$ in Tab. 7 is shown in Fig. 3. Compared to the conventional polarization sensor, our sparse polarization sensor obtains high-quality RGB images and polarization information, particularly for $r = 1/4$ and $1/16$, with excellent balance between the quality of each output. Further-

more, our compensation is effective in improving the quality of RGB images and polarization information for conventional polarization sensors.

5. Discussions and conclusion

We proposed a new sparse polarization sensor structure and network architecture that compensates for low-resolution polarization information to acquire high-quality RGB images and polarization information simultaneously. The results of this study are based on simulations. In the future, we intend to consider prototyping our spare polarization sensor to validate its practical application. Based on this research, we believe that polarization pixels could be incorporated into many cameras in the future.

Limitations: First, this study used a white color filter in the polarization pixel. Hence, we could not acquire wavelength-side polarization information. Developing a sparse polarization sensor structure and compensation network architecture to acquire wavelength information is a topic for future research. Second, cases where there is not a perfect correlation between the RGB image and polarization information, such as compensation in photoelastic measurements of transparent objects, are challenging to address with the current approach. Finally, our network is not suitable for hardware implementation owing to the large number of parameters (184.8M). We will work on making the network more compact in the future.

References

[1] Houdini. https://www.sidefx.com/products/houdini/. (Accessed on 07/05/2022). 2

[2] Unreal Engine Marketplace. https://www.unrealengine.com/marketplace/. (Accessed on 07/05/2022). 6

[3] G. A. Atkinson and E. R. Hancock. Multi-view surface reconstruction using polarization. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pages 309–316, 2005. 1

[4] G. A. Atkinson and E. R. Hancock. Recovery of surface orientation from diffuse polarization. IEEE Transactions on Image Processing (TIP), 2006. 1

[5] Yunhao Ba, Alex Gilbert, Franklin Wang, Jinfa Yang, Rui Chen, Yiqin Wang, Lei Yan, Boxin Shi, and Achuta Kadambi. Deep shape from polarization. In European Conference on Computer Vision (ECCV), pages 554–571. Springer, 2020. 6

[6] Seung-Hwan Baek and Felix Heide. Polarimetric spatio-temporal light transport probing. ACM Transactions on Graphics (TOG), 40(6):1–18, 2021. 1

[7] Seung-Hwan Baek, Daniel S. Jeon, Xin Tong, and Min H. Kim. Simultaneous acquisition of polarimetric SVBRDF and normals. ACM Transactions on Graphics (TOG), 37(6):268:1–15, 2018. 2

[8] Daniel Camozzato, Leandro Dihl, Ivan Silveira, Fernando Marson, and Soraia R Musse. Procedural floor plan generation from building sketches. The Visual Computer, 31(6):753–763, 2015. 6

[9] Chin-Cheng Chan, Shao-Kang Huang, and Homer H Chen. Enhancement of phase detection for autofocus. In IEEE International Conference on Image Processing (ICIP), pages 41–45. IEEE, 2017. 3

[10] Xinjing Cheng, Peng Wang, Chenye Guan, and Ruigang Yang. CSPN++: Learning context and resource aware convolutional spatial propagation networks for depth completion. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 34, pages 10615–10622, 2020. 3

[11] Zhaopeng Cui, Viktor Larsson, and Marc Pollefeys. Polarimetric relative pose estimation. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pages 2671–2680, 2019. 2

[12] Valentin Deschaintre, Yiming Lin, and Abhijeet Ghosh. Deep polarization imaging for 3D shape and SVBRDF acquisition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 15567–15576, 2021. 2

[13] Renwei Dian, Shutao Li, Anjing Guo, and Leyuan Fang. Deep hyperspectral image sharpening. IEEE Transactions on Neural Networks and Learning Systems (TNNLS), 29(11):5345–5355, 2018. 3

[14] Yoshiki Fukao, Ryo Kawahara, Shohei Nobuhara, and Ko Nishino. Polarimetric normal stereo. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 682–690, 2021. 2

[15] Abhijeet Ghosh, Graham Fyffe, Borom Tunwattanapong, Jay Busch, Xueming Yu, and Paul Debevec. Multiview face capture using polarized spherical gradient illumination. ACM Transactions on Graphics (TOG), 30:129, 12 2011. 2

[16] Viktor Gruev, Rob Perkins, and Timothy York. CCD polarization imaging sensor with aluminum nanowire optical filters. Optics Express, 2010. 1

[17] Jiaqi Gu, Zhiyu Xiang, Yuwen Ye, and Lingxuan Wang. DenseLiDAR: A real-time pseudo dense depth guided depth completion network. IEEE Robotics and Automation Letters, 6(2):1808–1815, 2021. 3

[18] Lingzhi He, Hongguang Zhu, Feng Li, Huihui Bai, Runmin Cong, Chunjie Zhang, Chunyu Lin, Meiqin Liu, and Yao Zhao. Towards fast and accurate real-world depth super-resolution: Benchmark dataset and baseline. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 9229–9238, 2021. 3, 7

[19] Mu Hu, Shuling Wang, Bin Li, Shiyu Ning, Li Fan, and Xiaojin Gong. Penet: Towards precise and efficient image guided depth completion. In IEEE International Conference on Robotics and Automation (ICRA), pages 13656–13662. IEEE, 2021. 3, 5

[20] Tomoki Ichikawa, Matthew Purri, Ryo Kawahara, Shohei Nobuhara, Kristin Dana, and Ko Nishino. Shape from sky: Polarimetric normal recovery under the sky. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 14832–14841, 2021. 2

[21] Andrey Ignatov, Nikolay Kobyshev, Radu Timofte, Kenneth Vanhoey, and Luc Van Gool. WESPE: Weakly supervised photo enhancer for digital cameras. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pages 691–700, 2018. 5

[22] Achuta Kadambi, Vage Taamazyan, Boxin Shi, and Ramesh Raskar. Polarized 3D: High-quality depth sensing with polarization cues. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pages 3370–3378, 2015. 2

[23] Agastya Kalra, Vage Taamazyan, Supreeth Krishna Rao, Kartik Venkataraman, Ramesh Raskar, and Achuta Kadambi. Deep polarization cues for transparent object segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 8602–8611, 2020. 1

[24] Beomjun Kim, Jean Ponce, and Bumsub Ham. Deformable kernel networks for joint image filtering. International Journal of Computer Vision (IJCV), 129(2):579–600, 2021. 3

[25] Masahiro Kobayashi, Michiko Johnson, Yoichi Wada, Hiromasa Tsuboi, Hideaki Takada, Kenji Togo, Takafumi Kishi, Hidekazu Takahashi, Takeshi Ichikawa, and Shunsuke Inoue. A low noise and high sensitivity image sensor with imaging and phase-difference detection af in all pixels. ITE Transactions on Media Technology and Applications (MTA), 4(2):123–128, 2016. 3

[26] Yuhi Kondo, Taishi Ono, Legong Sun, Yasutaka Hirasawa, and Jun Murayama. Accurate polarimetric BRDF for real polarization scene rendering. In European Conference on Computer Vision (ECCV), pages 220–236. Springer, 2020. 2

186

[27] Charis Lanaras, Emmanuel Baltsavias, and Konrad Schindler. Hyperspectral super-resolution by coupled spectral unmixing. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pages 3586–3594, 2015. 3

[28] Chenyang Lei, Xuhua Huang, Mengdi Zhang, Qiong Yan, Wenxiu Sun, and Qifeng Chen. Polarized reflection removal with perfect alignment in the wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 1750–1758, 2020. 2

[29] Chenyang Lei, Chenyang Qi, Jiaxin Xie, Na Fan, Vladlen Koltun, and Qifeng Chen. Shape from polarization for complex scenes in the wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 12632–12641, 2022. 6

[30] 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 European Conference on Computer Vision (ECCV), pages 457–473. Springer, 2020. 2

[31] Rui Li, Simeng Qiu, Guangming Zang, and Wolfgang Heidrich. Reflection separation via multi-bounce polarization state tracing. In European Conference on Computer Vision (ECCV), pages 781–796. Springer, 2020. 2

[32] Ruibo Li, Ke Xian, Chunhua Shen, Zhiguo Cao, Hao Lu, and Lingxiao Hang. Deep attention-based classification network for robust depth prediction. In Asian Conference on Computer Vision (ACCV), pages 663–678. Springer, 2018. 6

[33] Xiaobo Li, Haiyu Li, Yang Lin, Jianhua Guo, Jingyu Yang, Huanjing Yue, Kun Li, Chuan Li, Zhenzhou Cheng, Haofeng Hu, et al. Learning-based denoising for polarimetric images. Optics Express, 28(11):16309–16321, 2020. 3

[34] Yijun Li, Jia-Bin Huang, Narendra Ahuja, and Ming-Hsuan Yang. Joint image filtering with deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 41(8):1909–1923, 2019. 3

[35] Lina Liu, Xibin Song, Xiaoyang Lyu, Junwei Diao, Mengmeng Wang, Yong Liu, and Liangjun Zhang. FCFR-Net: Feature fusion based coarse-to-fine residual learning for depth completion. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 35, pages 2136–2144, 2021. 3

[36] Shumin Liu, Jiajia Chen, Yuan Xun, Xiaojin Zhao, and Chip-Hong Chang. A new polarization image demosaicking algorithm by exploiting inter-channel correlations with guided filtering. IEEE Transactions on Image Processing (TIP), 29:7076–7089, 2020. 3

[37] Youwei Lyu, Zhaopeng Cui, Si Li, Marc Pollefeys, and Boxin Shi. Reflection separation using a pair of unpolarized and polarized images. Advances in neural information processing systems, 32, 2019. 2

[38] 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 Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 12622–12631, 2022. 1

[39] Sofiane Mihoubi, Pierre-Jean Lapray, and Laurent Bigué. Survey of demosaicking methods for polarization filter array images. Sensors, 18(11):3688, 2018. 3

[40] Daisuke Miyazaki, Masataka Kagesawa, and Katsushi Ikeuchi. Transparent surface modeling from a pair of polarization images. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 26(1):73–82, 2004. 1

[41] Trung Ngo Thanh, Hajime Nagahara, and Rin-ichiro Taniguchi. Shape and light directions from shading and polarization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 2310–2318, 2015. 2

[42] Taishi Ono, Yuhi Kondo, Legong Sun, Teppei Kurita, and Yusuke Moriuchi. Degree-of-linear-polarization-based color constancy. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 19740–19749, June 2022. 6

[43] Jinsun Park, Kyungdon Joo, Zhe Hu, Chi-Kuei Liu, and In So Kweon. Non-local spatial propagation network for depth completion. In European Conference on Computer Vision (ECCV), pages 120–136. Springer, 2020. 3, 7

[44] Xuebin Qin, Zichen Zhang, Chenyang Huang, Masood Dehghan, Osmar R Zaiane, and Martin Jagersand. U2-Net: Going deeper with nested u-structure for salient object detection. Pattern Recognition, 106:107404, 2020. 2, 7

[45] Jiaxiong Qiu, Zhaopeng Cui, Yinda Zhang, Xingdi Zhang, Shuaicheng Liu, Bing Zeng, and Marc Pollefeys. DeepLiDAR: Deep surface normal guided depth prediction for outdoor scene from sparse lidar data and single color image. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 3313–3322, 2019. 3

[46] Simeng Qiu, Qiang Fu, Congli Wang, and Wolfgang Heidrich. Polarization demosaicing for monochrome and color polarization focal plane arrays. In Hans-Jörg Schulz, Matthias Teschner, and Michael Wimmer, editors, Vision, Modeling and Visualization. The Eurographics Association, 2019. 3

[47] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241. Springer, 2015. 7

[48] Wolfgang Sturzl. A lightweight single-camera polarization compass with covariance estimation. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pages 5353–5361, 2017. 2

[49] Kenichiro Tanaka, Yasuhiro Mukaigawa, and Achuta Kadambi. Polarized non-line-of-sight imaging. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 2136–2145, 2020. 1

[50] Jie Tang, Fei-Peng Tian, Wei Feng, Jian Li, and Ping Tan. Learning guided convolutional network for depth completion. IEEE Transactions on Image Processing (TIP), 30:1116–1129, 2020. 3, 7

[51] AB Tibbs, IM Daly, DR Bull, and NW Roberts. Noise creates polarization artefacts. Bioinspiration & biomimetics, 13(1):015005, 2017. 3

[52] Alexander B Tibbs, Ilse M Daly, Nicholas W Roberts, and David R Bull. Denoising imaging polarimetry by adapted BM3D method. Journal of the Optical Society of America A (JOSA A), 35(4):690–701, 2018. 3

[53] Juiwen Ting, Xuesong Wu, Kangkang Hu, and Hong Zhang. Deep snapshot HDR reconstruction based on the polarization camera. In IEEE International Conference on Image Processing (ICIP), pages 1769–1773. IEEE, 2021. 2

[54] Tali Treibitz and Yoav Y Schechner. Active polarization descattering. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 31(3):385–399, 2008. 1

[55] Masada Tzabari and Yoav Y Schechner. Polarized optical-flow gyroscope. In European Conference on Computer Vision (ECCV), pages 363–381. Springer, 2020. 2

[56] Tomohiro Yamazaki, Yasushi Maruyama, Yusuke Uesaka, Motoaki Nakamura, Yoshihisa Matoba, Takashi Terada, Kenta Komori, Yoshiyuki Ohba, Shinichi Arakawa, Yasutaka Hirasawa, et al. Four-directional pixel-wise polarization CMOS image sensor using air-gap wire grid on 2.5-μm back-illuminated pixels. In IEEE International Electron Devices Meeting (IEDM), pages 8–7. IEEE, 2016. 1

[57] Luwei Yang, Feitong Tan, Ao Li, Zhaopeng Cui, Yasutaka Furukawa, and Ping Tan. Polarimetric dense monocular slam. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 3857–3866, 2018. 2

[58] Yanchao Yang, Alex Wong, and Stefano Soatto. Dense Depth Posterior (DDP) from single image and sparse range. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 3353–3362, 2019. 3

[59] Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, and Nong Sang. Learning a discriminative feature network for semantic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 1857–1866, 2018. 6

[60] Wenfei Zhang, Jian Liang, Fei Xing, Zhongsheng Man, Xiaolu Ge, and Shenggui Fu. Polarimetric imaging method for target enhancement in haze based on polarimetric retrieval. Journal of Modern Optics, 66(11):1235–1243, 2019. 1

[61] Jinyu Zhao, Yusuke Monno, and Masatoshi Okutomi. Polarimetric multi-view inverse rendering. In European Conference on Computer Vision (ECCV), pages 85–102. Springer, 2020. 2

[62] Shanshan Zhao, Mingming Gong, Huan Fu, and Dacheng Tao. Adaptive context-aware multi-modal network for depth completion. IEEE Transactions on Image Processing (TIP), 30:5264–5276, 2021. 3

[63] Dizhong Zhu and William AP Smith. Depth from a polarisation + RGB stereo pair. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 7586–7595, 2019. 2

[64] Shihao Zou, Xinxin Zuo, Yiming Qian, Sen Wang, Chi Xu, Minglun Gong, and Li Cheng. 3D human shape reconstruction from a polarization image. In European Conference on Computer Vision (ECCV), pages 351–368. Springer, 2020. 2