A new polarization image demosaicking algorithm by exploiting inter-channel correlations with guided filtering

Liu, Shumin; Chen, Jiajia; Xun, Yuan; Zhao, Xiaojin; Chang, Chip-Hong

2020

Liu, S., Chen, J., Xun, Y., Zhao, X., & Chang, C.-H. (2020). A new polarization image demosaicking algorithm by exploiting inter-channel correlations with guided filtering. IEEE Transactions on Image Processing, 29, 7076-7089. doi:10.1109/TIP.2020.2998281

https://hdl.handle.net/10356/145824

https://doi.org/10.1109/TIP.2020.2998281

A New Polarization Image Demosaicking Algorithm by Exploiting Inter-Channel Correlations with Guided Filtering

Shumin Liu, Jiajia Chen, Yuan Xun, Xiaojin Zhao, *Member, IEEE*, and Chip-Hong Chang, *Fellow, IEEE*

***Abstract*—This paper presents a fast and effective polarization image demosaicking algorithm, which explores inter-channel dependency of Stokes parameters for the minimization of residual aliasing artifacts after cubic spline interpolation. A guided filtering approach is used for denoising. An optimization based on the confidence level of the aforementioned guided filtering, the correlations between the demosaicked image and input, as well as the total intensity, angle and degree of linear polarization, is constructed and solved with Newton's method. Experimental results demonstrate that the proposed algorithm can surpass the existing methods in terms of both objective root mean squared error and structural similarity index by at least 36.0% and 3.4%, respectively, and by close visual inspection of the clarity of objects in the angle and degree of linear polarization images. The proposed algorithm consists of only convolutions and element-wise operations, making it fast and parallelizable for efficient GPU acceleration. An image of size 512×612×4 can be processed within 10 s on i7-6700k CPU, and gains further 5 times speedup with M4000M GPU.**

***Index Terms*—Guided Filter, Image Demosaicking, Degree of Linear Polarization, Polarization Image, Image Denoising.**

I. Introduction

In addition to conventional intensity/color imaging, polarization imaging is another approach to overcome the limitations of object detection in natural scenes by intensity-based and color-based image processing. Inspired by some animals' instinct of polarization vision [1], polarization imaging has been applied to biomedical image analysis [2]-[3], 3-dimensional object recognition [4] and material classification [5]-[6], to name a few. Particularly, surfaces of artificial targets reflect and radiate electromagnetic waves with a significantly larger degree of linear polarization of reflection than surfaces of live objects from the same angle of view [7]. Thus, polarization images can improve the contrast ratio of artificial target and live objects from their distinction in both the angle and degree of linear polarization images [8]-[9].

Among different polarization image sensor implementations, division of focal plane (DoFP) polarization imager has become the mainstream. It enables the polarization sensing system to be integrated on a single chip, leading to high-resolution imaging at significantly reduced manufacturing cost [10]-[11]. To visualize the polarization state, a single image sensor is overlaid with an array of subwavelength spaced patterned wire grids to winnow specific polarization components. This is similar to the well-known mosaicked color filter array used widely in color imaging (e.g., Bayer pattern) [12]. With the rapid development of nano-scale semiconductor fabrication process, the pixelated micro-polarizing elements can be directly patterned and fabricated on top of the commercial CCD/CMOS image sensor's photodiode array to extract and measure the different polarization components of the incident light. However, DoFP polarization makes a trade-off between sensing speed (i.e. frame rate) and spatial resolution as different polarization components from the incident light are sampled simultaneously in a single frame. As a result, "polarization demosaicking" is a crucial imaging pipeline in DoFP polarization images [13]-[15]. It is used to reconstruct a high-quality full-resolution image from the down-sampled polarization sub-images. As each pixel element of the image sensor can detect only one channel of raw image data, the data of the missing channels will have to be predicted from the spatially shifted sensor data [16]. Since spatial-limited signals can never be band-limited, aliasing is inevitable when a full-resolution image is reconstructed by sampling the raw image data, resulting in zipper, fringes and other visual artifacts. Thus, a good demosaicking algorithm is required to render a high-quality image by reducing the residue artifacts upon interpolation. Although demosaicking algorithms for color images have been intensively studied [14]-[20], they cannot be directly applied to polarization images due to two fundamental differences. First of all, the channel configurations in the filter arrays for color and polarization image sensors are different. In the case of color images, demosaicking is performed among three wavelength channels, e.g., Bayer pattern with red, blue and green channels. As human perception is more sensitive to green, the weight of green sensors is doubled in the color filter

Manuscript received 19 Aug. 2019, revised 15 Jan. and 20 Mar. 2020, accepted 21 May 2020. This work was supported by Nanjing University of Aeronautics and Astronautics under Grant 56YAH18043 and PAC19009, Nanjing, China, and the Singapore Ministry of Education AcRF Tier 1 Grant MOE2018-T1-001-131, RG87/18-(S). (*Corresponding Author: Jiajia Chen*)

The author Shumin Liu is with the Pillar of Engineering Product Development, Singapore University of Technology and Design, Singapore.

The authors Jiajia Chen and Yuan Xun are with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, China.

The author Xiaojin Zhao is with the College of Electronics and Information Engineering, Shenzhen University, China.

The author Chip-Hong Chang is the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore.

array, as shown in Fig. 1(a) [12]. On the other hand, the four channels (0°, 45°, 90° and 135°) for polarization imaging are equally important, thus the number of sensors in different channels are identical, as shown in Fig. 1(b) [21]. The second difference lies in the correlation between different channels. For color images, RGB channels are less dependent on each other. In contrast, the four channels of polarization images are intricately interdependent [22]. Taking the 0° channel as an example, it is more correlated to the 45° and 135° channels than the 90° channel [22].

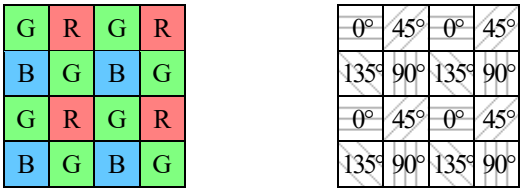

Fig. 1 Digital image filter arrays for (a) color and (b) polarization imaging.

Polarization image demosaicking can be considered as a super-resolution problem, where the input image in each channel is enlarged by a factor of two in both horizontal and vertical directions. In order to improve the accuracy, various interpolation algorithms such as bilinear, bicubic, spline and gradient-based methods have been proposed [23]-[29]. In general, the effect of these interpolations is analogous to low-pass filtering, which smoothens the transition of adjacent pixel values by filling in the missing values. Although a low root mean squared error (RMSE) can be achieved from the interpolated outputs, the reconstructed images are often blurry with poorly-preserved high-frequency information. To restore the high-frequency information from the residuals, the residual of interpolation was introduced in [30]. Recently, machine learning methods have been used for interpolation. These methods include sparse representation [31] that makes use of sparsity and non-local self-similarity priors to regularize interpolation and convolutional neural networks (CNN) [32] as an end-to-end framework to generate the interpolated Stoke images directly. Although deep learning (DL) methods can automatically deduce and optimally tune the features, the desired outcomes are achieved with very long and costly training. Besides requiring expensive computing resources, a huge amount of labeled data is also needed to train the CNN. Although in principle DL-based approach is adaptable to different applications and data types, the portability to new applications is constrained by the difficulty to comprehend the output based on mere learning without the classifiers, and the lack of standard theory in selecting the right DL topology, learning method and model hyperparameter settings. Based on our latest literature review, DL algorithm for polarization image demosaicking has not been reported. One main reason is the lack of labeled polarization image datasets for training the CNN. Besides interpolation, the output RMSE can be reduced by denoising algorithms such as Gaussian modeling [33] or block-matching and 3D filtering (BM3D) [34]. The high-frequency information can be replenished from fine features and textures that repeat themselves at multiple locations in an image [34].

The above-mentioned algorithms consider merely the intra-channel dependency for interpolation. In this paper, the inter-channel dependency between different polarization angles and Stokes parameters are also explored, which leads to the following three main contributions. Firstly, we propose a fast and accurate DoFP interpolation method exploiting the inter-channel correlations. Secondly, we propose a combination denoising scheme guided filters (GF) to effective denoise the interpolated images, and a novel method to estimate the confidence level of each predicted pixel value through GF for further denoising. Thirdly, a holistic optimization is formulated with only two optimizing parameters, which can be efficiently solved by Newton's method. This optimization is guided by AoLP, DoLP, $S0$ formulated in (1)-(5) and three sets of confidence levels computed through GF. The quality of the reconstructed polarization image is significantly enhanced with the proposed inter-channel correlation algorithm.

$$S0 = \frac{X_{\cdot,1} + X_{\cdot,2} + X_{\cdot,3} + X_{\cdot,4}}{2} \tag{1}$$

$$S1 = X_{\cdot,1} - X_{\cdot,4} \tag{2}$$

$$S2 = X_{\cdot,2} - X_{\cdot,3} \tag{3}$$

$$\mathrm{AoLP} = \frac{1}{2}\mathbf{arctan}\left(\frac{S2}{S1}\right) \tag{4}$$

$$\mathrm{DoLP} = \frac{\sqrt{S1^2 + S2^2}}{S0} \tag{5}$$

The rest of this paper is organized as follows. Section II introduces the basic concept and the proposed variations of guided filtering. Section III presents the problem formulation and the details of the proposed algorithm. In Section IV, the performance and computational efficiency of the proposed algorithm are evaluated, analyzed and compared with related state-of-the-art methods. The paper is concluded in Section V.

II. Guided Filtering and the Proposed Variations

A. The Original Guided Filtering

GF is a denoising algorithm that is well-known for its edge-preserving property and efficiency [35]. There are two inputs to GF, which are a target image p to be denoised and a guidance image I. The inputs of GF are separated into a sequence of overlapping square windows with a stride of 1. Each window w_k of size $(2r + 1) \times (2r + 1)$ centered at pixel k is processed independently. The output q is a linear transformation of I, which can be expressed as:

$$q_i = a_k I_i + b_k \quad \forall i \in w_k \tag{6}$$

where q_i and I_i are the values of q and I, respectively at pixel i, and a_k and b_k are the regression coefficients for the window w_k.

There are $(2r + 1) \times (2r + 1)$ pairs of (q_i, I_i) for w_k. The values of a_k and b_k can be optimized by minimizing the "unexplained variation", which is defined as the sum of squared differences between input p and output q. The objective function of GF can be expressed mathematically [36] as:

$$\arg\min_{a_k,b_k}\sum_{i\in\omega_k}\left(\left(a_kI_i+b_k-p_i\right)^2+\epsilon a_k^2\right). \tag{7}$$

The optimization of the unexplained variation of (7) can be seen as a "straight line of best fit" problem, where a_k is the gradient and b_k is the y-axis interception. The regularization parameter ϵ reduces the variances of the estimated regression parameters, and avoids a zero-value denominator in (8). The optimization problem in (7) has a closed-form solution expressed in (8) and (9),

$$a_k=\frac{\frac{1}{|w_k|}\sum_{i\in w_k}\left(I_ip_i-\mu(I_k)\mu(p_k)\right)}{\delta(I_k)^2+\epsilon} \tag{8}$$

$$b_k=\mu(p_k)-a_k\mu(I_k) \tag{9}$$

where $\mu(I_k)$ and $\mu(p_k)$ are respectively the means of I and p, calculated within w_k, $\delta(I_k)$ is the standard deviation of I within w_k, and $|w_k|$ is the number of pixels in ω_k.

From (7) to (9), we can compute the optimized a_k and b_k, and then substitute them into (6) to obtain the GF output. Since the inputs of GF are separated into overlapping square windows, and each window has $(2r+1)\times(2r+1)$ outputs, there will be $(2r+1)\times(2r+1)$ outputs generated at each pixel position (except for the boundary pixels). The blue outlined 3×3 windows for $r=1$ are shown in Fig. 2. The final output $\bar{q}_i$ at pixel i is calculated by taking the average value of all q_i containing the pixel i from different windows. Mathematically, $\bar{q}_i$ can be calculated by (10).

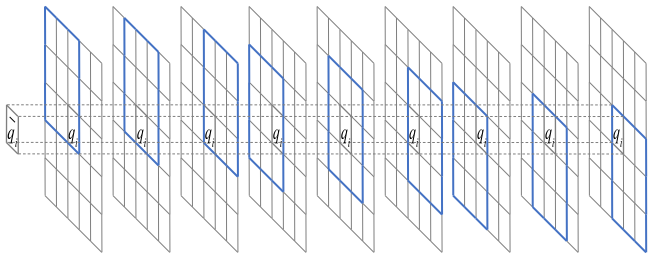

Fig. 2 Calculation of $\bar{q}_i$ from q_i of all 3×3 windows containing the pixel i.

$$\bar{q}_i=\frac{1}{|w_k|}\sum_{k\forall i\in w_k}\left(a_kI_i+b_k\right)=\bar{a}_iI_i+\bar{b}_i \tag{10}$$

where

$$\bar{a}_i=\frac{1}{|w_k|}\sum_{k\forall i\in w_k}a_k;\ \ \bar{b}_i=\frac{1}{|w_k|}\sum_{k\forall i\in w_k}b_k\ . \tag{11}$$

B. Confidence-aware Guided Filtering

The original GF gives the output $\bar{q}$. In polarization image demosaicking, the Stokes parameters are inter-related. Thus, they can be leveraged for mutual denoising. By varying the guidance image I, multiple $\bar{q}$'s can be generated from the same target image p. To fully utilize different $\bar{q}$'s, a confidence level L is proposed to estimate the cleanliness of each $\bar{q}$. The confident level L_i of q_i at pixel i is defined as the reciprocal of the unexplained variation for w_k, i.e.,

$$L_i=1/U_k\quad\forall i\in w_k \tag{12}$$

Similar to (6), there are $(2r+1)\times(2r+1)$ L_i, and U_k is a constant for w_k. U_k is optimized in a similar way as (7), i.e.,

$$U_k=\min_{a_k,b_k}\sum_{i\in\omega_k}\left(\left(a_kI_i+b_k-p_i\right)^2+\epsilon a_k^2\right) \tag{13}$$

which has the following closed-form solution for the linear regression [36]:

$$U_k=|w_k|\left(\delta(p_k)^2-a_k^2\delta(I_k)^2\right) \tag{14}$$

where a_k, $\delta(I_k)$ and $|w_k|$ are defined in (8). $\delta(p_k)$ is the standard deviation of p within the window w_k. It is the only new function requiring additional computation on top of (8). Similar to Fig. 2 and (10), the confidence level $\bar{L}_i$ of $\bar{q}_i$ at pixel i, is calculated by averaging all L_i from different windows having the pixel i.

$$\bar{L}_i=\frac{1}{|w_k|}\sum_{k\forall i\in w_k}1/U_k \tag{15}$$

$\bar{L}_i$ is a measure of how confident the value of $\bar{q}_i$ is close to the ground truth. The higher the value of $\bar{L}_i$, the less the deviation of $\bar{q}_i$ from its true value. The confidence level approaches positive infinity when the unexplained variation approaches zero. This happens when p and I are identical.

The entire denoised image $\bar{q}$ and the confidence level $\bar{L}$ can be generated by calculating $\bar{q}_i$ and $\bar{L}_i$ at each pixel from (10) and (15). For simplicity, this calculation can be expressed as a single function **GF** in (16), where the input arguments to the function are a target image p to be denoised and a guidance image I, and the outputs of the function are $\bar{q}$ and $\bar{L}$.

$$\left(\bar{q},\bar{L}\right)=\mathbf{GF}\left(p,I\right) \tag{16}$$

C. Angular Guided Filtering

GF is originally designed to denoise pixel values in Cartesian coordinate system. However, some pixel values, such as AoLP, are represented in polar coordinate system. To enable GF on AoLP, a new angular guided filtering (AGF) is proposed. Similar to GF, the inputs to AGF are the target image p and guidance image I; The outputs are the denoised image $\bar{q}$ and its confidence level $\bar{L}$ expressed in (17).

$$\left(\bar{q},\bar{L}\right)=\mathbf{AGF}\left(p,I\right) \tag{17}$$

Three modifications are required to change GF to AGF:

1. The range of AoLP is expanded to [−π, π]. As AoLP ranges from [−π/4, π/4] originally, the range can be expanded by simply multiplying AoLP by 4.
2. AoLP is converted to the complex Cartesian system by Euler's formula. It suffices to apply AGF only on AoLP with guidance image $S0$. In other words, among the two inputs p and I, only the conversion for p is required. Steps 1 and 2 are expressed as a single function (18), where **GF**(·), **exp**(·) and j are the original GF function, exponential function and unit

imaginary number respectively.

$$\left(\overline{q}_{\text{Cartesian}}, \overline{L}_{\text{Cartesian}}\right) = \mathbf{GF}\left(\mathbf{exp}\left(p \times 4 \times j\right), I\right) \tag{18}$$

3. The output of GF, $\overline{q}_{\text{Cartesian}}$, is converted back to the polar system by (19), where **img**(·) and **real**(·) are the imaginary and real parts, respectively of a complex number, and **arctan**(·) is the inverse tangent function. The output $\overline{q}_{\text{polar}}$ is rescaled back to $\overline{q}$ in the range [−π/4, π/4] in (20). Similar operations are performed on $\overline{L}_{\text{Cartesian}}$.

$$\overline{q}_{\text{polar}} = \mathbf{arctan}\left(\frac{\mathbf{img}\left(\overline{q}_{\text{Cartesian}}\right)}{\mathbf{real}\left(\overline{q}_{\text{Cartesian}}\right)}\right) \tag{19}$$

$$\overline{q} = \overline{q}_{\text{polar}} / 4 \tag{20}$$

With the original GF, the proposed confidence level and the proposed AGF for AoLP presented in this section, a new demosaicking algorithm is proposed, where both the RMSE and SSIM of the demosaicked image are improved by the denoised $\overline{q}$. The regularization parameter of demosaicking optimization is set by confidence $\overline{L}$. The denoising and regression optimization with AoLP can be processed with the help of AGF. The details will be presented in the next section.

III. The Proposed Algorithm

The proposed algorithm consists of three main functions as depicted in Fig. 3, which are interpolation, denoising with I and denoising with AoLP/DoLP. $I \in \mathbb{Z}^{N\times M\times 4}$ is the polarized input image and $Y \in \mathbb{Z}^{2N\times 2M\times 4}$ is the output. N and M are the length and width, respectively of the input image in terms of the number of pixels. The first function makes use of the correlations between four input channels to generate a preliminary demosaicked image; the second function makes use of the correlations between demosaicked image, input image and $S0$ to denoise the demosaicked image; the third function makes use of the correlations between $S0$, AoLP and DoLP to further denoise the demosaicked image. In short, all these steps are designed based on one overarching principle of incorporating as many correlations between polarization channels as possible to improve the demosaicking quality, hence the name inter-channel correlations (ICC) of our proposed algorithm. These three functions are elaborated as follows.

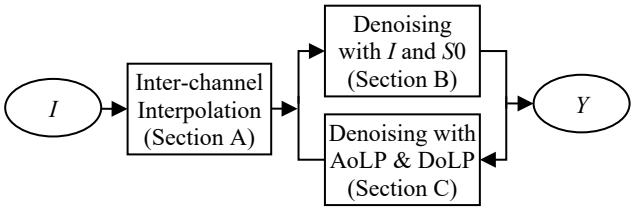

Fig. 3 A high-level overview of the proposed algorithm.

A. Inter-channel Interpolation

Interpolation is a fundamental preliminary step in demosaicking. As a key input to subsequent denoising and optimization processes, the interpolated images could significantly affect the overall demosaicking quality. There exist many types of single image interpolation methods, such as linear, nearest pixel value, cubic spline [37] and makima [41]. Unfortunately, they cannot be easily adapted to polarization demosaicking, as the important inter-channel correlation is not considered. The interpolation of [42] is founded on the assumption that the luminance channel is highly correlated with all three color channels. The same assumption is also valid for polarization demosaicking as the total intensity ($S0$) is highly correlated with all four polarization channels. The design of our interpolation filter is inspired by this observation.

Three representative filters shown in Fig. 4 are required for color demosaicking in [42], but only two representative filters, as shown in Fig. 5(a) and (b), are required for polarization demosaicking. Besides, the filters in Fig. 5(b) and (c) are the transpose of each other. The filter in Fig. 4(a) has the same pattern as that in Fig. 5(a), thus the coefficients can be directly transplanted. The filter in Fig. 4(b) has a similar pattern as that in Fig. 5(b), with four extra "−1"s in the diagonal directions of the central "5". Thus, these "−1"s are added into the coefficients on the fringe to obtain the filter shown in Fig. 5(b). The filter in Fig. 4(c) is useless for polarization images and can be omitted.

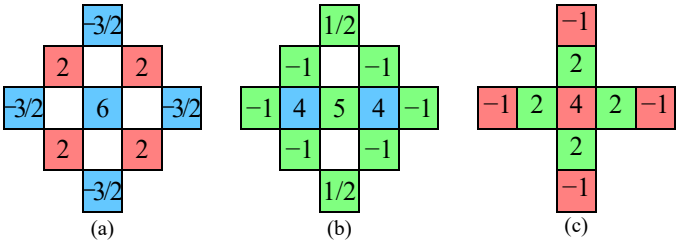

Fig. 4 Three representative color demosaicking filters in [42], where (a), (b) and (c) are the demosaicking of red pixel in blue position, blue pixel in green position and green pixel in red position, respectively.

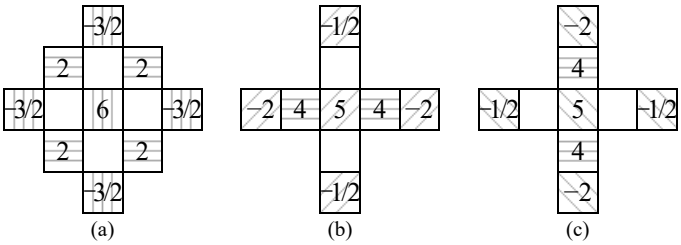

Fig. 5 The proposed demosaicking filters for polarization images, where (a), (b) and (c) are the demosaicking of 0° pixel in 90° position, 45° position and 135° position, respectively.

When the DoFP input image is convoluted by the three proposed filters in Fig. 5, the three missing channels of DoFP input can be predicted. For simplicity, this set of convolutions is expressed as a single line function **InterChannel** in (21), where $I \in \mathbb{Z}^{N\times M\times 4}$ and $X \in \mathbb{Z}^{2N\times 2M\times 4}$ are the input and preliminary interpolated image respectively.

$$X = \mathbf{InterChannel}\left(I\right) \tag{21}$$

B. Denoising by GF with Input Image and S0

The output of inter-channel interpolation X is denoised by GF. As mentioned in Section II, a guidance image, that is less noisy and closely correlated to X is required. Both input image I

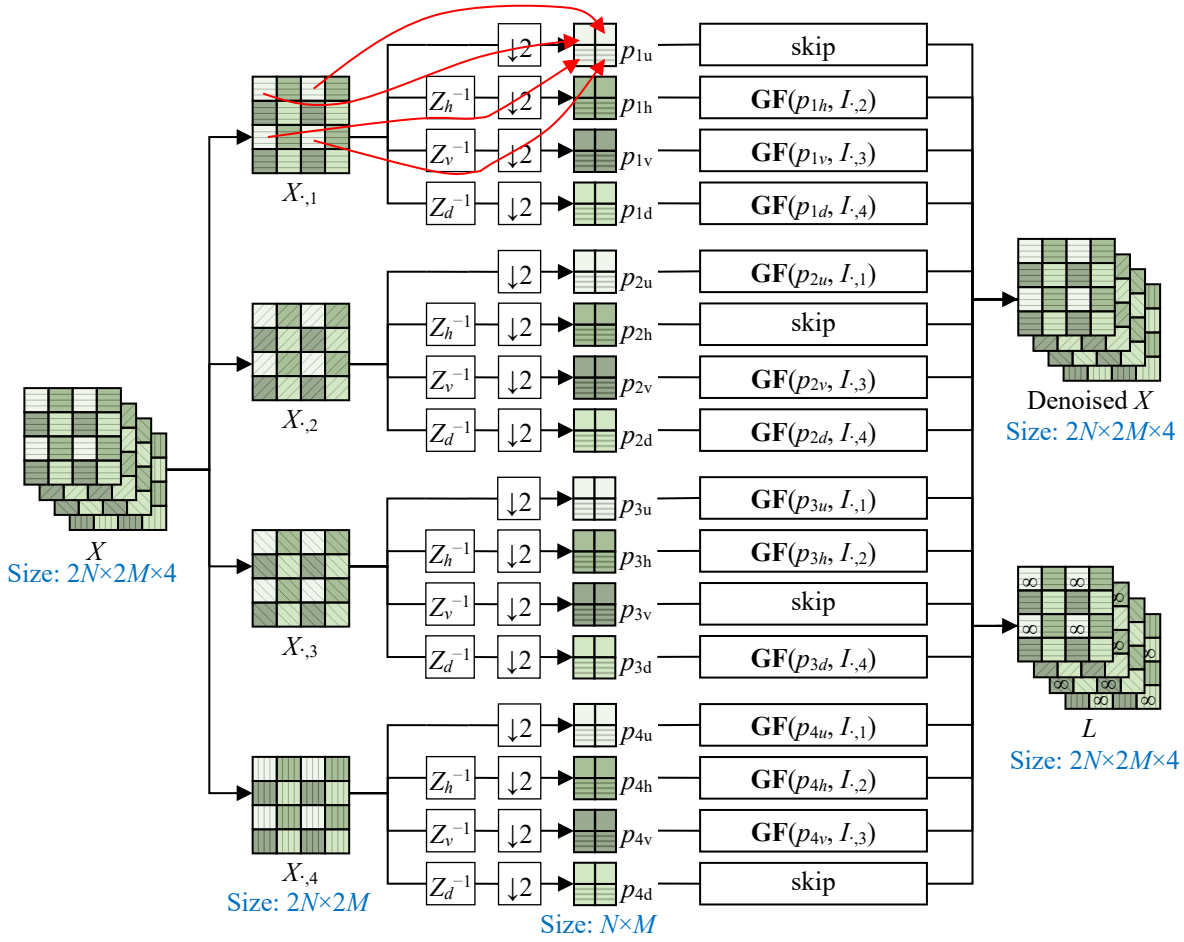

Fig. 6 Denoising by 12 GFs with the input image as guidance image.

and total intensity $S0 \in \mathbb{Z}^{2N\times2M}$ fulfill the above two requirements. Among I and $S0$, I is less noisy but $S0$ has a higher resolution. As a result, $S0$ complements I to achieve a high-quality denoising of X. In what follows, GF denoising by exploiting the inter-channel correlations between X, I and $S0$ is presented, where X is denoised by two GFs guided by I and $S0$, respectively, and the two denoised Xs are fused with the confidence levels generated by GFs.

1) Denoising by GF with Input Image

There are four channels each in X and I. Consider all possible combinations of pairing a channel of X with a channel of I, 16 GFs will be needed to fully denoise X. Of which, four GFs are trivial as their denoised images are identical to the guidance images. Since the output is the same as the input for these four GFs, they are redundant and can be omitted.

The operations of the remaining 12 GFs are shown in Fig. 6. The denoising image X is first divided into four channels, which are $X_{\cdot,j} \in \mathbb{Z}^{2N\times2M}$ with $j \in [1, 4]$. For each $X_{\cdot,j}$, it is further divided into four sub-images, p_{ju}, p_{jh}, p_{jv} and p_{jd}. The sub-image p_{ju} is generated directly by down-sampling $X_{\cdot,j}$ by a factor of two (denoted by ↓2 in Fig. 6) using the simple nearest-neighbor method [43]. The other three sub-images, p_{jh}, p_{jv} and p_{jd}, are produced after shifting $X_{\cdot,j}$ by one pixel in the horizontal, vertical and diagonal directions (denoted by Z_h^{-1}, Z_v^{-1} and Z_d^{-1} in Fig. 6), respectively, followed by factor of two down-sampling. The GF operation of (16) is then applied to these sub-images. The guidance images of GF are chosen based on the position of p and $I_{\cdot,j}$. As shown in Fig. 7, the sub-images, p_{ju}, p_{jh}, p_{jv} and p_{jd}, are respectively guided by $I_{\cdot,1}$, $I_{\cdot,2}$, $I_{\cdot,3}$ and $I_{\cdot,4}$ that are located at the same pixel positions.

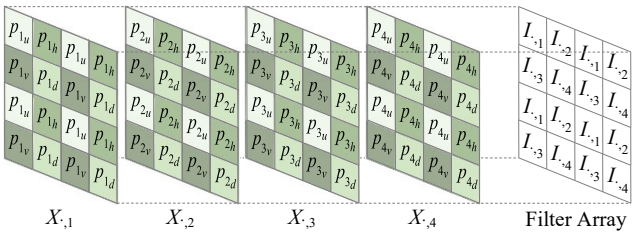

Fig. 7 The correspondence of pixel locations of sub-images, p_{ju}, p_{jh}, p_{jv} and p_{jd}, on $X_{\cdot,j}$, and input image channels, $I_{\cdot,1}$, $I_{\cdot,2}$, $I_{\cdot,3}$ and $I_{\cdot,4}$, on polarization filter array, where the background colors of X correspond to those in Fig. 6.

Two outputs, namely the denoised X_I and its confidence level $LX_I \in \mathbb{R}^{2N\times2M\times4}$, are generated after applying the twelve GF functions. For those pixels that have skipped the four trivial GF functions, their confidence levels are recorded as positive infinity. This process is described by the following single line function in the pseudo-code of our algorithm.

$$(X_I, LX_I) = \textbf{DenoiseByI}(X, I) \tag{22}$$

2) Denoising by GF with S0

The noise in each of the four X channels of $S0$ is assumed to

have the same normal distribution with zero mean and variance σ^2. Without loss of generality, the noise of $S0$, $n(S0)$, can be considered as an average of the noise components $n(X_{\cdot,j})$ contributed by the random variables $X_{\cdot,j}$ for all channels j, $j \in [1, 4]$. Based on the statistical property of the sum and scaling of independent normal distributions, the mean and variance of $n(S0) = [n(X_{\cdot,1}) + n(X_{\cdot,2}) + n(X_{\cdot,3}) + n(X_{\cdot,4})]/4$ are zero and $\sigma^2/4$, respectively. Since the standard deviation of the noise of $S0$ is $\sigma/2$, which is only half that of X, $S0$ can be used by GF to denoise X. The four channels of X are respectively denoised by GF with $S0$. The denoised X_{S0} and its confidence level $LX_{S0} \in \mathbb{R}^{2N\times 2M\times 4}$ are output. This process is described by the following single line function.

$$\left(X_{S0}, LX_{S0}\right) = \mathbf{DenoiseByS0}\left(X, S0\right) \tag{23}$$

3) Fusion

X_I and X_{S0} are fused by the following pixel-wise comparison between LX_I and LX_{S0}:

$$X = \begin{cases} X_I & LX_I > LX_{S0} \\ X_{S0} & \text{otherwise} \end{cases}. \tag{24}$$

The pixels with higher confidence levels are selected. The same operation is performed for the confidence level:

$$LX = \begin{cases} LX_I & LX_I > LX_{S0} \\ LX_{S0} & \text{otherwise} \end{cases}. \tag{25}$$

C. Optimization with GF-denoised AoLP and DoLP

Different Stokes representations can be generated from X, which include total intensity ($S0$), angle of linear polarization (AoLP) and degree of linear polarization (DoLP), where $S0$, $S1$, $S2$, AoLP, DoLP $\in \mathbb{Z}^{2N\times 2M}$. The AoLP and DoLP are first denoised by $S0$ using GFs. The operations are expressed as (26) and (27), where LA and $LD \in \mathbb{Z}^{2N\times 2M}$ are the confidence level of the denoised AoLP and DoLP, respectively.

$$\left(\text{AoLP}, LA\right) = \mathbf{AGF}\left(\text{AoLP}, S0\right) \tag{26}$$

$$\left(\text{DoLP}, LD\right) = \mathbf{GF}\left(\text{DoLP}, S0\right) \tag{27}$$

The denoised AoLP and DoLP are used to adjust X by solving the regression problem in (28). All variables in (28) and (29) are scalar, and the optimizations are done pixel-wise. For each variable, the first subscript represents the pixel position while the second represents the index of the channel. A variable of only one channel is represented with only one subscript,

$$\min_{\substack{X_{i,1}, X_{i,2} \\ X_{i,3}, X_{i,4}}} \left\{ \begin{array}{c} LA_i \times \left(\frac{1}{2}\mathbf{arctan}\left(\frac{X_{i,2} - X_{i,3}}{X_{i,1} - X_{i,4}} \right) - \text{AoLP}_i \right)^2 + \\ LD_i \times \left(\frac{\sqrt{\left(X_{i,1} - X_{i,4}\right)^2 + \left(X_{i,2} - X_{i,3}\right)^2}}{S0_i} - \text{DoLP}_i \right)^2 \\ + \epsilon \sum_{j=1}^{4} LX_{i,j} \left(X_{i,j} - X'_{i,j} \right)^2 \end{array} \right\} \tag{28}$$

where $X_{i,j} \in \mathbb{Z}$ and $0 \le X_{i,j} \le 255$. $X_{i,j}$ and $L_{i,j}$ are the values of X and LX at the i^{th} pixel of the j^{th} channel calculated by (22). $S0_i$, AoLP_i and DoLP_i, LA_i and LD_i are the values of $S0$, AoLP and DoLP, LA and LD, respectively at the i^{th} pixel. Note that $X'_{i,j}$ are the original values of the arguments $X_{i,j}$ for $j = 1$, 2, 3 and 4, after (24), before they are updated to the new values by solving the minimization problem. ϵ is a positive regularization parameter with a very small magnitude. The minimization in (28) aims to adjust the values of X such that they can produce AoLP and DoLP nearest to the denoised values (represented by pixel values $AoLP_i$ and $DoLP_i$, respectively). There are three terms in (28). The first and second terms are regularized by LA and LD, respectively. A larger LA value indicates a greater confidence in $AoLP_i$ than $DoLP_i$. Thus, $\left(1/2 \cdot \mathbf{arctan}\left(S2_i/S1_i\right) - \text{AoLP}_i\right)^2$ of the first term has a smaller magnitude than $\left(\sqrt{\left(S1_i\right)^2 + \left(S2_i\right)^2}/S0_i - \text{DoLP}_i\right)^2$ of the second term. The last term aims to constrain the change in X to a minimum magnitude.

There are four variables in (28) to be optimized. It is difficult and time-consuming to solve them independently. To simplify (28), we replace $X_{i,1} - X_{i,4}$ with $S1_i$ and $X_{i,2} - X_{i,3}$ with $S2_i$ such that,

$$\begin{aligned} \min_{\substack{S1_i \\ S2_i}} & \left\{ \begin{array}{l} LA_i \times \left(\frac{1}{2}\mathbf{arctan}\left(\frac{S2_i}{S1_i} \right) - \text{AoLP}_i \right)^2 + \\ LD_i \times \left(\frac{\sqrt{\left(S1_i\right)^2 + \left(S2_i\right)^2}}{S0_i} - \text{DoLP}_i \right)^2 \end{array} \right\} \\ \text{s.t.} \quad & \mathbf{sgn}\left(X_{i,1} - X'_{i,1}\right) = -\mathbf{sgn}\left(X_{i,4} - X'_{i,4}\right) \\ & \mathbf{sgn}\left(X_{i,2} - X'_{i,2}\right) = -\mathbf{sgn}\left(X_{i,3} - X'_{i,3}\right) \\ & LX_{i,4}\left(X_{i,1} - X'_{i,1}\right) = -LX_{i,1}\left(X_{i,4} - X'_{i,4}\right) \\ & LX_{i,3}\left(X_{i,2} - X'_{i,2}\right) = -LX_{i,2}\left(X_{i,3} - X'_{i,3}\right) \end{aligned} \tag{29}$$

where $S1_i$ and $S2_i$ are the values of $S1$ and $S2$, respectively at the i^{th} pixel. In (28), $X_{i,1}$ and $X_{i,4}$ are always calculated in pairs. Thus, the difference between them can be replaced by $S1_i$ without losing any accuracy. $X_{i,1}$ and $X_{i,4}$ must change in opposite directions in order to minimize $(X_{i,1} - X'_{i,1})$ and $(X_{i,4} - X'_{i,4})$ for a given $S1_i$. This leads to the first constraint. After $S1_i$ is optimized, it is straightforward to calculate the changes in $X_{i,1}$ and $X_{i,4}$ as they must be inversely proportional to $LX_{i,1}$ and $LX_{i,4}$ in order to minimize $LX_{i,1} \times (X_{i,1} - X'_{i,1})^2 + LX_{i,4} \times (X_{i,4} - X'_{i,4})^2$. This leads to the third constraint. The same is true for $X_{i,2}$ and $X_{i,3}$, which leads to the second and fourth constraints in (29).

Equation (29) can be directly optimized by Newton's method. The optimization starts with the original values of $S1$ and $S2$, this reduces the probability of overfitting. The smallest step size is limited to 0.1 to accelerate the optimization. This will not appreciably affect the accuracy as the output pixels are stored as integers. With this step size limit, over 99% of the pixel values of $S1$ and $S2$ will converge to the optimal values by the 10^{th} iterations. Further improvement after the 10^{th} iteration, if any, is trivial based on our experimentation. After the optimal values of $S1$ and $S2$ are found, the optimal value of

X can be calculated by,

$$\begin{aligned} X_{i,1} &= X'_{i,1} + LX_{i,4} \Big/ \left(LX_{i,1} + LX_{i,4}\right) \times \left(S1_i - S1'_i\right) \\ X_{i,2} &= X'_{i,2} + LX_{i,3} \Big/ \left(LX_{i,2} + LX_{i,3}\right) \times \left(S2_i - S2'_i\right) \\ X_{i,3} &= X'_{i,3} - LX_{i,2} \Big/ \left(LX_{i,2} + LX_{i,3}\right) \times \left(S2_i - S2'_i\right) \\ X_{i,4} &= X'_{i,4} - LX_{i,1} \Big/ \left(LX_{i,1} + LX_{i,4}\right) \times \left(S1_i - S1'_i\right) \end{aligned} \tag{30}$$

where $S1'_i$ and $S2'_i$ are the original values of $S1_i$ and $S2_i$, respectively. All pixels of X are constrained to the range [0, 255] to produce the final output. The complete algorithm is described by the pseudo-code in Fig. 8. For ease of reference, the calling functions are annotated by the section and equation numbers where they are introduced.

```
Function ICC(I)
    //Section III.A
    X = InterChannel(I)   //Eqn. (21)
    for i = 1 to iteration
        //Section III.B
        (X_I, LX_I) = DenoisingByI(X, I)        //Eqn. (22)
        (X_S0, LX_S0) = DenoisingByS0(X, I)     //Eqn. (23)
        (X, LX) = GetXandLX(X_I, LX_I, X_S0, LX_S0)     //Eqn. (24)-(25)
        //Section III.C
        (S0, S1, S2, DoLP, AoLP) = GetStokes(X)     //Eqn. (1)-(5)
        (AoLP, LA) = AGF(AoLP, S0)                  //Eqn. (26)
        (DoLP, LD) = GF(DoLP, S0)                   //Eqn. (27)
        (S1, S2) = Optimize(S1, S2, AoLP, DoLP, LA, LD)   //Eqn. (29)
        X = Update(S1, S2, LX)                      //Eqn. (30)
    Return Min(Max(X, 0), 255)
```

Fig. 8 Proposed inter-channel correlation (ICC) polarization image demosaicking algorithm.

IV. Experimental Results and Discussions

In this section, the performance of the proposed demosaicking method, ICC, is evaluated together with four state-of-the-art polarization demosaicking methods, leveraging on intensity correlation (IC) method [24], sparse representation (SR) [31], Newton's polynomial (NP) [38], and block-matching and 3-D filtering denoising (BM3D) denoising [34] respectively. Ten polarization images with input size 256×306×4 (output size 512×612×4) are used as the benchmark images [39], which are listed in Fig. 9.

Fig. 9 Benchmark images.

We use FLIR BFS-U3-51S5P-C camera to capture the DoFP images. The DoFP images are separated into four polarization channels, followed by spline interpolation to estimate the missing pixels. The interpolated images are then downsized by a factor of four in both horizontal and vertical directions to produce the synthetic ground truth. In other words, one pixel in the synthetic ground truth is generated by averaging 16 pixels in the interpolated image to preserve the accuracy of pixel values in the ground truth. At last, the synthetic inputs to ICC is generated by down-sampling. The size of image at each stage is illustrated in the flowchart of Fig. 10. This is a simple way to synthesize polarization image pairs from DoFP images. The synthetic ground truth and synthetic input to ICC may be used as training pairs to deep learning networks to supplement the scarcity of real labeled data.

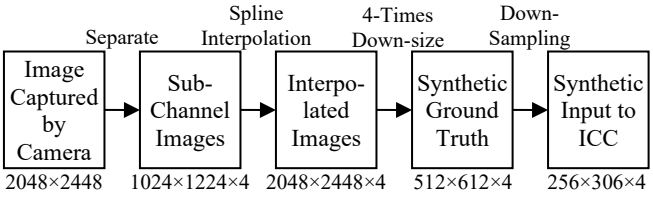

Fig. 10 The procedure for synthesizing the input and ground truth images.

To evaluate the demosaicking quality, we measure the RMSE and SSIM of the outputs X, $S0$, AoLP and DoLP against the ground truth. Similar to Section II.C, the values of AoLP are stored in the polar system. Thus, modifications on RMSE and SSIM are required for the evaluation of AoLP. For angular SSIM (ASSIM), the same modification as AGF is made, except that the output is rescaled to [0, 1], as shown in (31) and (32),

$$output_{\text{Cartesian}} = \mathbf{SSIM}\left(\mathbf{exp}\left(\text{AoLP}\times 4\times i\right), \mathbf{exp}\left(\widehat{\text{AoLP}}\times 4\times i\right)\right) \tag{31}$$

$$output = \mathbf{arctan}\left(\frac{\mathbf{img}\left(output_{\text{Cartesian}}\right)}{\mathbf{real}\left(output_{\text{Cartesian}}\right)}\right) / \pi \tag{32}$$

where AoLP and $\widehat{\text{AoLP}}$ are the angle of linear polarization of the output and the ground truth, respectively.

For angular RMSE (ARMSE), the following modification is performed,

$$\text{ARMSE}\left(\begin{matrix}\text{AoLP},\\ \widehat{\text{AoLP}}\end{matrix}\right) = \sqrt{\frac{1}{4NM}\sum_{\forall i}\left(\mathbf{min}\left(\begin{matrix}\text{AoLP}_i - \widehat{\text{AoLP}}_i,\\ \pi/2 - \text{AoLP}_i + \widehat{\text{AoLP}}_i\end{matrix}\right)\right)^2} \tag{33}$$

where pixel-level intensity is denoted by indexing the variable with subscript i. N and M are the length and width of the AoLP, respectively. Similar to RMSR and SSIM, a lower ARMSE and a higher ASSIM imply a better demosaicking quality. For simplicity, the ARMSE and ASSIM of AoLP are also referred to as RMSE and SSIM, respectively. To compare each method effectively across X, $S0$, AoLP and DoLP, the normalized RMSE is computed. The RMSE for each Stokes parameter is first normalized by the RMSE for the same parameter of ICC algorithm before they are normalized across all Stokes parameters. The same operations are performed to obtain the normalized SSIM.

The experiments in this section are organized as follows. Firstly, the parameters used in the ICC algorithm are evaluated. Secondly, the demosaicking performance of ICC is compared with four competing methods. Finally, the complexity and processing speed of ICC are analyzed.

TABLE I
COMPARISON OF DEMOSAICKING QUALITY IN RMSE AND SSIM WITH/WITHOUT PROPOSED TECHNIQUES

Technique	RMSE					SSIM				
	X	*S*0	AoLP	DoLP	Normalized	*X*	*S*0	AoLP	DoLP	Normalized
Linear Interpolation	3.5052	6.1657	0.2612	0.0243	1.3004	0.9471	0.9206	0.7229	0.8304	0.9898
Cubic Interpolation	3.4336	5.9763	0.2633	0.0254	1.2946	0.9476	0.9209	0.7198	0.8284	0.9883
Makima Interpolation	3.3421	5.8195	0.2601	0.0242	1.2579	0.9492	0.9227	0.7246	0.8329	0.9923
Spline Interpolation	3.5151	6.1118	0.2654	0.0279	1.3400	0.9451	0.9167	0.7151	0.8253	0.9841
Inter-Channel Interpolation	**2.3662**	**3.6786**	**0.2560**	**0.0237**	**1**	**0.9633**	**0.9311**	**0.7252**	**0.8384**	**1**
Denoising by GF with *S*0	2.6457	4.3666	0.2576	0.0243	1.0838	0.9583	0.9276	0.7152	**0.8386**	0.9944
Denoising by GF with *I*	2.5367	**3.5157**	0.2587	0.0258	1.0314	0.9589	0.9200	**0.7291**	0.8242	0.9930
Denoising by GF with *I* and *S*0	**2.3662**	3.6786	**0.2560**	**0.0237**	**1**	**0.9633**	**0.9311**	0.7252	0.8384	**1**
Optimization w/o Confidence	2.4672	3.8305	**0.2556**	0.0243	1.0274	0.9599	0.9282	0.7023	0.8237	0.9861
Optimization with Confidence	**2.3662**	**3.6786**	0.2560	**0.0237**	**1**	**0.9633**	**0.9311**	**0.7252**	**0.8384**	**1**

A. Effectiveness of the Proposed Techniques

In this section, the effectiveness of proposed inter-channel interpolation, denoising by GF with *I* and *S*0, confidence level for optimization, and iterative denoising are investigated. The experimental results are summarized in Table I. The contribution of each technique used in ICC is evaluated by comparing its RSME and SSIM with those achieved by replacing it with an alternative technique listed in the first column of Table I. For example, the first row of Table I displays the RSME and SSIM of ICC when the proposed inter-channel interpolation is replaced with linear interpolation. For interpolation, the results in the first five rows show that the proposed inter-channel interpolation is better than all the other four methods of interpolation for all Stokes parameters, with at least 25.7% and 0.7% improvements in normalized RMSE and SSIM, respectively. The last two rows indicate that the proposed confidence level improves the demosaicking quality in all parameters except the RMSE of AoLP, with a negligible difference of only 0.15%. The overall improvements are 2.7% and 1.3% in normalized RMSE and SSIM, respectively. For denoising, the results in the sixth to eighth rows show that most parameters benefit from the GF denoising by a combination of *I* and *S*0, with at least 3.1% and 0.5% improvements in RMSE and SSIM, respectively.

To investigate the effectiveness of iterative denoising and determine the number of iterations, RMSE and SSIM are plotted against the numbers of iterations in Fig. 11. The mean RMSE and SSIM of the ten benchmark images in [39] are used for this evaluation. All RMSEs are normalized by the raw RMSE measured directly after the cubic spline interpolation and before any GF operation, while SSIM is plotted without any post-processing. A half iteration refers to the completion of denoising with input, i.e., after the 12 GFs in Section III.B, and a full iteration refers to the completion of regression optimization with the denoised AoLP/DoLP in Section III.C. The effectiveness of denoising with input can be evaluated by comparing the RMSE and SSIM of the output image between iteration k and $k+0.5$. Meanwhile, the performance of the optimization with AoLP/DoLP can be evaluated by comparing between iterations $k+0.5$ and $k+1$. From Fig. 11, it can be observed that the RMSE and SSIM for all the Stokes images improve rapidly in the first half iteration and then gradually converge to relatively steady values. After the first half iteration, the RMSE can still be reduced by both techniques. On the other hand, as the name suggests, the improvements in SSIM of AoLP/DoLP are due mainly to the optimization with AoLP/DoLP. These two operations complement each other to improve the quality of the demosaicked images. However, RSME and SSIM do not converge at the same iteration. Moreover, GF introduces some minor artifacts, such as "detail halos" [44], in each iteration. These artifacts impact SSIM more than RMSE. Because of the artifacts of GF, it can be seen from Fig. 11 that although the RMSE curves are flattened at the fourth iteration, the improvement in the overall SSIM ceases after the second iteration. Thus, a good balance between the artifact and denoising power of GF is achieved by terminating the algorithm after two iterations for benchmark images [39].

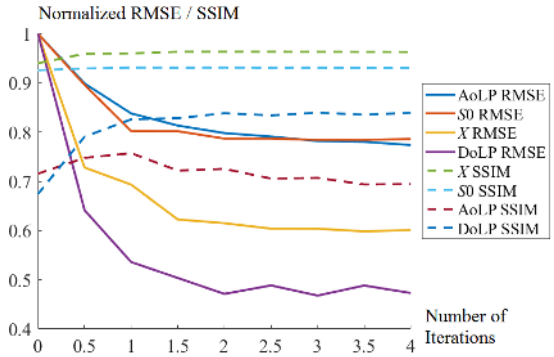

Fig. 11 Normalized RMSE and SSIM after every half iteration of the ICC algorithm.

B. Performance Evaluation of ICC

In order to perform a comprehensive evaluation, this section is divided into three sub-sections to test the performance of ICC by clean images in [39], noisy images in [39], and other images in [31] and [38].

1) Clean Benchmark Images in [39]

The demosaicking quality of ICC is compared with four state-of-the-art algorithms, both objectively in terms of RMSE/SSIM and subjectively by detailed inspection of a benchmark image [39]. The mean RMSE of the ten benchmark images is summarized in the third to seventh columns of Table II. Among these five demosaicking methods, the ICC algorithm

TABLE II
COMPARISON OF DEMOSAICKING QUALITY IN RMSE AND SSIM OF DIFFERENT ALGORITHMS

Input Images	Method	RMSE					SSIM				
		X	*S*0	AoLP	DoLP	Normalized	*X*	*S*0	AoLP	DoLP	Normalized
Images in [39]	IC [24]	5.7458	8.2828	0.3406	0.0658	2.1970	0.8797	0.8677	0.7039	0.6168	0.8879
	SR [31]	4.5548	6.8298	0.3066	0.0500	1.7725	0.9140	0.8893	**0.7343**	0.7099	0.9408
	NP [38]	3.2101	4.4976	0.3014	0.0399	1.3603	0.9468	0.9190	0.7283	0.7430	0.9651
	BM3D [34]	6.6742	11.9280	0.6427	0.0505	2.6765	0.8040	0.6947	0.4850	0.7455	0.7847
	ICC	**2.3662**	**3.6786**	**0.2560**	**0.0237**	**1**	**0.9633**	**0.9311**	0.7252	**0.8384**	**1**
Images in [39] Magnitude of Noise: 1	IC [24]	5.8130	8.3338	0.3574	0.0675	2.1877	0.8531	0.8302	0.6656	0.5948	0.8723
	SR [31]	4.5992	6.8793	0.3167	0.0509	1.7562	0.8966	0.8587	0.7104	0.6951	0.9375
	NP [38]	3.3152	4.7073	0.3126	0.0407	1.3732	0.9214	0.8570	0.6984	0.7315	0.9503
	BM3D [34]	6.6826	11.9380	0.6394	0.0507	2.6089	0.7639	0.6260	0.6308	0.7386	0.8216
	ICC	**2.4378**	**3.8531**	**0.2578**	**0.0240**	**1**	**0.9459**	**0.8829**	**0.7109**	**0.8371**	**1**
Images in [39] Magnitude of Noise: 2	IC [24]	6.0068	8.4768	0.3754	0.0732	2.1795	0.8157	0.7780	0.6160	0.5494	0.8487
	SR [31]	4.7369	7.0349	0.3342	0.0536	1.7329	0.8725	0.8126	0.6621	0.6601	0.9247
	NP [38]	3.6085	5.2801	0.3294	0.0432	1.4120	0.8852	0.7792	0.6507	0.7046	0.9271
	BM3D [34]	6.6995	11.9440	0.6321	0.0511	2.4634	0.7594	0.6233	0.6010	0.7314	0.8351
	ICC	**2.6365**	**4.3223**	**0.2600**	**0.0241**	**1**	**0.9225**	**0.8147**	**0.6868**	**0.8341**	**1**
Images in [39] Magnitude of Noise: 5	IC [24]	7.2024	9.4100	0.4099	0.1053	2.1798	0.6919	0.6502	0.5439	0.3953	0.7936
	SR [31]	5.6524	8.0392	0.3760	0.0718	1.6864	0.7740	**0.6883**	0.5735	0.5099	0.8798
	NP [38]	5.1661	8.1245	0.3717	0.0627	1.5696	0.7611	0.6174	0.5687	0.5913	0.8720
	BM3D [34]	6.8178	11.9790	0.6056	0.0544	1.9657	0.7408	0.6176	0.5544	0.6921	0.8913
	ICC	**3.7209**	**6.6982**	**0.2691**	**0.0273**	**1**	**0.8328**	0.6566	**0.6269**	**0.8136**	**1**
Images in [39] Magnitude of Noise: 10	IC [24]	10.3160	12.0770	0.4362	0.1733	2.2758	0.5302	0.5203	0.5142	0.2224	0.7440
	SR [31]	8.3688	**10.7990**	0.4160	0.1229	1.7979	0.6107	0.5532	0.5238	0.2931	0.8169
	NP [38]	8.6183	14.0950	0.4067	0.1103	1.7817	0.5927	0.4662	0.5258	0.4154	0.8091
	BM3D [34]	7.3773	12.1930	0.5643	0.0704	1.5490	0.6831	**0.5878**	0.5232	0.5695	0.9520
	ICC	**6.1817**	11.7164	**0.2845**	**0.0356**	**1**	**0.6936**	0.5064	**0.5794**	**0.7503**	**1**
Image "Buildings" in [31]	IC [24]	1.5633	2.2211	0.0658	0.0097	1.5094	0.9834	0.8919	0.8648	0.9424	0.9575
	SR [31]	1.1106	1.6893	0.0439	**0.0069**	1.0738	0.9894	0.9112	0.9157	**0.9686**	0.9846
	NP [38]	2.1057	2.9222	0.0716	0.0188	2.0663	0.9753	0.9327	0.8875	0.9432	0.9724
	BM3D [34]	2.4655	4.6612	0.7613	0.0091	6.8618	0.9561	0.6284	0.8464	0.9625	0.8824
	ICC	**1.0168**	**1.3612**	**0.0370**	0.0089	**1**	**0.9929**	**0.9624**	**0.9237**	0.9651	**1**
Images in [45]	IC [24]	2.0133	2.2711	0.2392	**0.0624**	1.0163	0.9187	0.8403	0.8855	**0.7267**	0.9957
	SR [31]	1.9333	2.0894	0.2535	0.0817	1.0751	0.8425	0.7431	0.8303	0.5870	0.8823
	NP [38]	2.3918	2.8771	0.2320	0.0939	1.2452	0.8998	0.8351	0.8687	0.5805	0.9329
	BM3D [34]	2.5269	3.4387	0.6544	0.0979	1.8089	0.7415	0.4194	0.7619	0.4346	0.6887
	ICC	**1.9570**	**2.2180**	**0.2257**	0.0655	**1**	**0.9275**	**0.8658**	**0.8882**	0.7090	**1**

performs the best in terms of RMSE for all Stokes parameters. Specifically, ICC can reduce the RMSE of existing algorithms further by at least 35.6%, 22.2%, 17.7% and 68.3% for *X*, *S*0, AoLP and DoLP, respectively. The reduction of RMSE in DoLP by ICC over other methods is most prominent, which demonstrates the effectiveness of the denoising process proposed in Section III.C. The improvement in RMSE is the least for AoLP, as GF is a linear denoiser, while AoLP involves trigonometry function. From the seventh column in Table II, our proposed ICC algorithm could reduce the overall RMSE by at least 36.0%. ICC does not result in over-smoothing, as evinced by the SSIM scores listed in the third to seventh columns of Table II. It is observed that ICC improves the SSIM of *X* and *S*0 by at least 1.7% and 1.3% respectively compared with the existing algorithms. For the similar reason that GF is a linear denoiser, ICC improves the SSIM of DoLP by 12.4% but the AoLP score is slightly reduced by no more than 1.3% in comparison with the best competing method. Nonetheless, the normalized SSIM of ICC is improved by at least 3.4%, as listed in the last column of Table II. Based on the objective RMSE and SSIM scores, it can be concluded that the demosaicking quality of ICC is the best among the five methods.

For a subjective evaluation of the demosaicking quality, both the AoLP and DoLP images of the benchmark image "carpark" generated by the five algorithms are compared in Fig. 12. It is evident that the AoLP and DoLP images produced by the ICC algorithm in (f) and (l) are closest to the ground truth images in (a) and (g), respectively. The DoLP produced by the competitor methods in (h) to (k) tends to have larger pixel values, which makes DoLP brighter than the ground truth. This inaccuracy can also be observed in AoLP, where the competitor methods tend to generate noisy AOLP in (b) to (e), making it more "colorful" than the ground truth. The inaccuracy of AoLP and DoLP by the competitor methods is due to the fact that they tend to insert artificial information on the edges of the features. On the other hand, the proposed ICC method focuses on retrieving the correct pixel values. For example, the road markings in the ground truth DoLP in (g) is black in color, which are accurately demosaicked by ICC in (l). In contrast, those markings in (h) to (k) by IC, SR NP and BM3D respec-

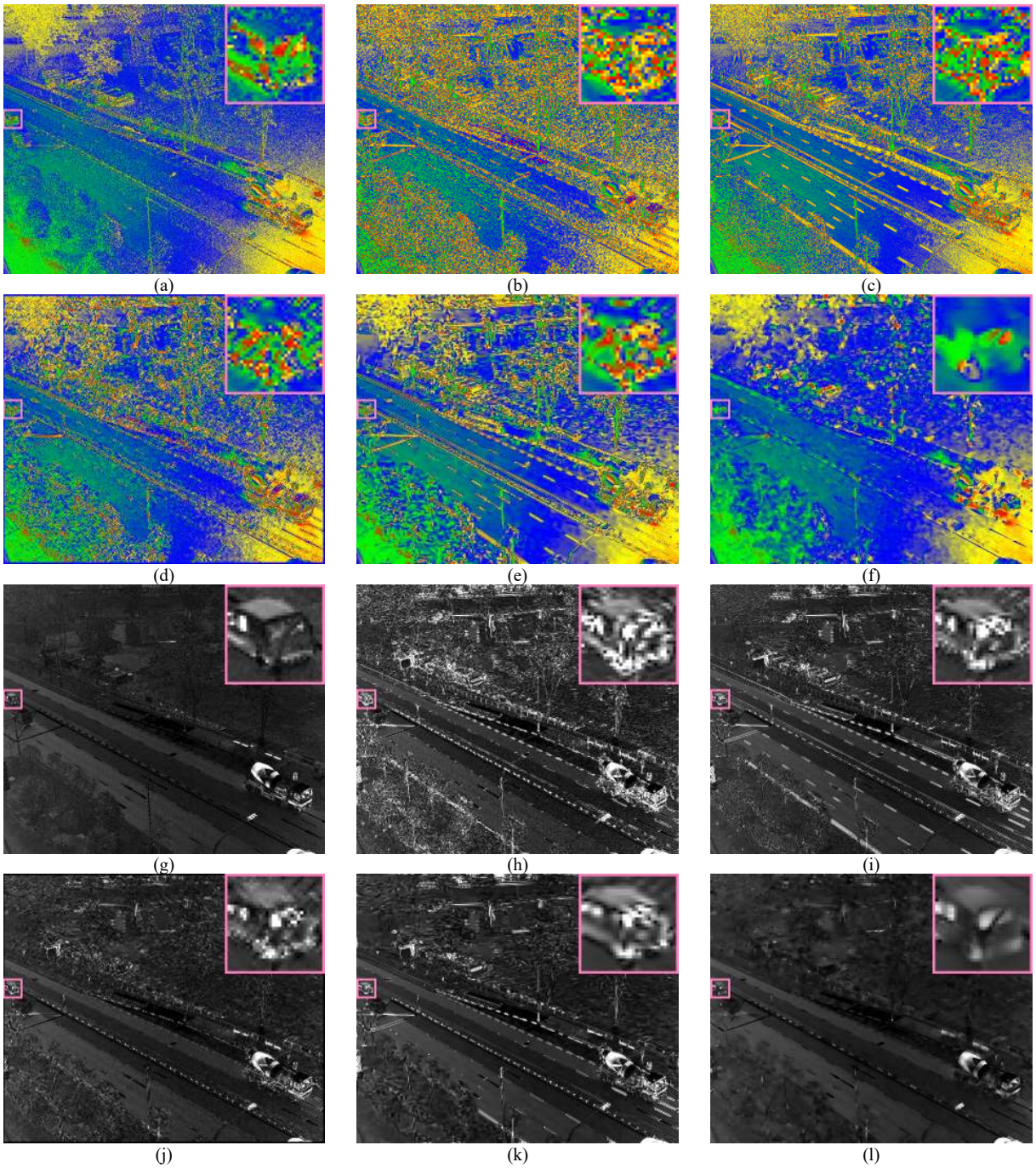

Fig. 12 Subjective evaluation of demosaicking quality of benchmark image "carpark", where (a) is the AoLP ground truth, (b), (c), (d), (e) and (f) are the AoLP images produced by IC, SR, NP, BM3D and the proposed ICC, respectively, where red, yellow, blue and green represent polarization angle 0, π/8, π/4 and -π/8 respectively; (g) is the DoLP ground truth, (h), (i), (j), (k) and (l) are the DoLP images produced by IC, SR, NP, BM3D and the proposed ICC, respectively.

tively are much brighter, as the edges of the markings are over-emphasized and recolored into white. Although emphasizing the edge contrast may sharpen certain features of the DoLP images, they cause more loss of fidelity by introducing noises and unwanted information that are detrimental to accurate and correct information retrieval. A similar observation can be drawn from the trees at the bottom left corner, where the trees look sharper in (i) of SR and (j) of NP. This is partially because ICC is less capable of demosaicking fine textures. As the input has a lower resolution than DoLP, those fine features of size less than two pixels in DoLP may be ignored during denoising. For this reason, fine textures are noisy in (f). Nonetheless, these trees produced by SR and NP are noisier, as they look more reddish and brighter in AoLP and DoLP, respectively. Apart from the accuracy of independent features, ICC achieves the best demosaicking when two objects intercept. This can be seen when a street lamp partially blocks a car, as magnified on the top-right corner of each image. Both the lamp and car in (f) and (l) produced by ICC are clearer, and more distinguishable than those produced by the competing methods. In conclusion, the proposed ICC algorithm generates the most accurate AoLP and DoLP among all methods.

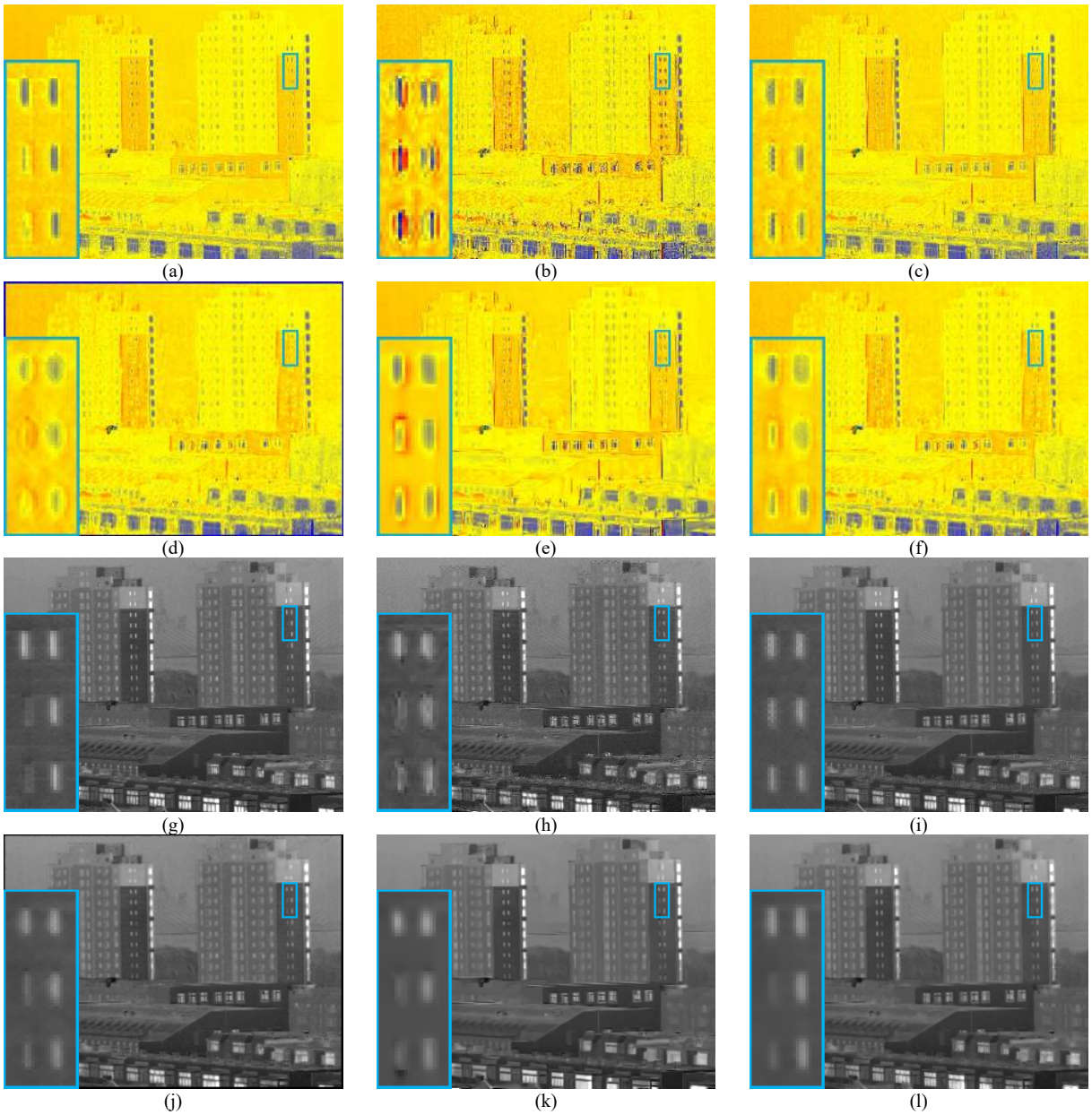

Fig. 13 Subjective evaluation of demosaicking quality of benchmark image "buildings", where (a) is the AoLP ground truth, (b), (c), (d), (e) and (f) are the AoLP images produced by IC, SR, NP, BM3D and the proposed ICC, respectively, where red, yellow, blue and green represent polarization angle 0, π/8, π/4 and -π/8 respectively; (g) is the DoLP ground truth, (h), (i), (j), (k) and (l) are the DoLP images produced by IC, SR, NP, BM3D and the proposed ICC, respectively.

2) Noisy Benchmark Images in [39]

As real input DoFP image may be noisy, to evaluate the effect of noise, four sets of experiments are conducted with noisy inputs. The root mean square magnitudes of noises are set to be 1, 2, 5 and 10, respectively. Their results are shown in the sixth to 25th row of Table II. It can be seen that ICC scores the best in RMSE in most cases and in SSIM for the majority of the cases. This is due to the strong denoising power of GF inherited into ICC and this is maximally exploited in each step. In terms of normalized RMSE and SSIM, ICC always scores the best among the five algorithms, with an improvement of at least 37.3% and 4.8% respectively compared with the best competing method. In conclusion, as observed from the results of these four sets of experiments, ICC can still perform even with noisy DoFP images.

3) Other Benchmark Images in [31] and [38]

Apart from [39], we have also applied the five demosaicking algorithms to a set of real polarization images provided by SR [31]. The RMSE and SSIM scores are summarized in the 26th to 30th rows of Table II. From the objective evaluations of RMSE and SSIM, a similar conclusion can be drawn on the best performance of ICC, as it scores the best for six out of eight metrics. In addition, ICC improves the normalized RMSE and SSIM by 7.3% and 1.5%, respectively. The AoLP and DoLP images of all five algorithms are displayed in Fig. 13. It can be seen that the images produced by ICC algorithm in (f) and (l) are the cleanest images, and the content is nearest to the ground truth. A zoomed-in view of six windows is given on the bottom left corner of each image in Fig. 13, where two of the windows are smaller in size and look faint in both the AoLP and DoLP

ground truth images. The subtle features are much better preserved by ICC, as shown in (f) and (l). We have also tested on 10 multispectral image sets in [45], where the RMSE and SSIM scores are summarized in the last five rows of Table II. In this test, ICC scores the best in six out of eight metrics, resulting in at least 1.6% and 0.4% improvements on RMSE and SSIM, respectively. This again supports our findings that the demosaicking quality of ICC is the best among the five methods in comparison.

4) Performance with Deep Learning Denoising

To the best of our knowledge, there is no DL method dedicated to polarization demosaicking in the literature, making it impossible to compare a non-DL method with a DL method for the same application and dataset. It is, however, feasible to evaluate if the proposed method will benefit from DL by considering a denoising CNN (DCNN) [46] and a super-resolution CNN (SRCNN) [47] as a plug-in in the demosaicking task. Similar to the frameworks proposed in [46] and [47], the input images are inter-channel interpolated before they are fed into the respective DL network. The performances of ICC, and ICC with DCCN and SRCNN plug-ins are compared in Table III. The results show that as direct plug-in to ICC, DL-based denoising method actually degrades its demosaicking quality, as evinced by the poorer RMSE and SSIM values.

TABLE III
COMPARISON BETWEEN ICC AND DEEP LEARNING METHODS

Methods	RMSE				
	X	*S*0	AoLP	DoLP	Normalized
ICC w DCNN [46]	3.8444	4.7428	0.6304	0.0497	1.8686
ICC w SRCNN [47]	3.8486	4.6755	0.6291	0.0503	1.8697
ICC	**2.3662**	**3.6786**	**0.2560**	**0.0237**	**1**
Methods	SSIM				
	X	*S*0	AoLP	DoLP	Normalized
ICC w DCNN [46]	0.9382	0.9227	0.7165	0.6777	0.9404
ICC w SRCNN [47]	0.9399	0.9258	0.7154	0.6740	0.9401
ICC	**0.9633**	**0.9311**	**0.7252**	**0.8384**	**1**

C. Algorithmic Complexity and Processing time

Except for cubic spline interpolation and GF, all other operations of the ICC algorithm are pixel-wise operations. Spline and GF require convolutions with a fixed-size filter. Hence, the overall complexity of ICC is $O(NM)$, where N and M are the length and width of the input image I. This makes ICC effective for high-resolution images. For a fair comparison of their execution time, all five algorithms are implemented in Matlab 2017b and run on the same PC with i7-6700k CPU and 32GB RAM.

TABLE IV
COMPARISON OF CPU TIME (IN SECONDS) OF DIFFERENT ALGORITHMS

Method	Input Image Sizes		
	128×153×4	256×306×4	512×612×4
IC [24]	0.0365	0.2541	1.2133
SR [31]	192.66	817.28	3555.3
NP [38]	0.0406	0.4237	1.9790
BM3D [34]	2.7144	11.356	48.625
ICC	0.3650	2.7412	10.9834
ICC on GPU	0.2883	1.0217	1.8080

The runtime is reported in Table IV. The processing speed of ICC is two orders of magnitude faster than SR. It is noteworthy that SR requires 23.3GB memory to processing the input of size 512×512×4. SR is slow and resource-intensive because it is based on machine learning. Training and matching of dictionary are required in order to learn the non-local redundancy of images. Although ICC is slower than IC and NP, this small processing time overhead is more than paid off by the significantly improved output quality, as the outputs generated by IC and NP are 119.7% and 36.0%, respectively noisier than ICC. As all operations are carried out pixel-wise except for convolutions, ICC can be easily accelerated with GPU. By converting all matrices to gpuArray and pixel-wise operations to arrayfun in Matlab, the processing time can be reduced by 21.0%, 62.7% and 83.5% with an M4000M GPU for small, medium and large size images, respectively as depicted in the last row of Table IV. This proves that ICC is beneficial to parallelism.

V. CONCLUSION

This paper presents a novel polarization image demosaicking algorithm with output quality exceeding the state-of-the-art methods reported in recent years. After interpolation, two denoising processes are carried out iteratively by exploiting the correlations between channels and Stokes parameters of the interpolated image, which converges with small RMSE and large SSIM scores rapidly in two iterations. The proposed algorithm improves the normalized RMSE and SSIM of four recently proposed algorithms by at least 36.0% and 3.4% respectively. Moreover, the proposed algorithm has low computational complexity. Its pixel-wise and convolution operations are highly parallelizable to aid GPU acceleration. For large polarization images, its processing speed can be accelerated by GPU to run 5 times faster than on CPU. In the future, we will explore dedicated deep network models for polarization demosaicking based on the findings and synthetic inputs made possible through this research.

REFERENCES

[1] N. W. Roberts, M. J. How, M. L. Porter, S. E. Temple, R. L. Caldwell, S. B. Powell, V. Gruev, N. J. Marshall and T. W. Cronin, "Animal polarization imaging and implications for optical processing," *Proceedings of the IEEE*, vol. 102, no. 10, pp. 1427-1434, Oct. 2014.

[2] B. Kunnen, C. Macdonald, A. Doronin, S. Jacques, M. Eccles and I. Meglinski, "Application of circularly polarized light for non-invasive diagnosis of cancerous tissues and turbid tissue-like scattering media," *Journal of Biophotonics*, vol. 8, no. 4, pp. 317-323, Apr. 2015.

[3] M. Zhang, X. Wu, N. Cui, N. Engheta and J. Van der Spiegel, "Bioinspired focal-plane polarization image sensor design: From application to implementation," *Proceedings of the IEEE*, vol. 102, no. 10, pp. 1435-1449, Oct. 2014.

[4] Z. Chen, X. Wang and R. Liang, "Snapshot phase shift fringe projection 3D surface measurement," *Optics Express*, vol. 23, no. 2, pp. 667-673, Jan. 2015.

[5] T. V. T. Krishna, C. D. Creusere and D. G. Voelz, "Passive polarimetric imagery-based material classification robust to illumination source position and viewpoint," *IEEE Transactions on Image Processing*, vol. 20, no. 1, pp. 288-292, Jan. 2011.

[6] M. Sarkar, D. San Segundo Bello, C. Van Hoof and A. Theuwissen, "Integrated polarization analyzing CMOS image sensor for material classification," *IEEE Sensors Journal*, vol. 11, no. 8, pp. 1692-1703, Aug. 2011.

[7] E. M. Rudd, M. Günther and T. E. Boult, "PARAPH: presentation attack rejection by analyzing polarization hypotheses," *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Las Vegas, US, Jun. 2016, pp. 171-178.

[8] M. Garcia, C. Edmiston, R. Marinov, A. Vail and V. Gruev, "Bioinspired color-polarization imager for real-time in situ imaging," *Optica*, vol. 4, no. 10, pp. 1263-1271, Oct. 2017.

[9] N. M. Garcia, I. de Erausquin, C. Edmiston and V. Gruev, "Surface normal reconstruction using circularly polarized light," *Optics Express*, vol. 23, no. 11, pp. 14391-14406, Jun. 2015.

[10] S. Li, W. Ye, H. Liang, X. Pan, X. Lou and X. Zhao, "K-SVD based denoising algorithm for DoFP polarization image sensors," *IEEE International Symposium on Circuits and Systems*, Florence, Italy, May 2018, pp. 1-5.

[11] M. Garcia, S. Gao, C. Edmiston, T. York and V. Gruev, "A 1300 × 800, 700 mW, 30 fps spectral polarization imager," *IEEE International Symposium on Circuits and Systems*, Lisbon, Portugal, May 2015, pp. 1106-1109.

[12] B. E. Bayer, "Color imaging array," U.S. Patent 3 971 065, Jul. 1976.

[13] C. Bai, J. Li, Z. Lin, J. Yu and Y. Chen, "Penrose demosaicking," *IEEE Transactions on Image Processing*, vol. 24, no. 5, pp. 1672-1684, May 2015.

[14] D. Kiku, Y. Monno, M. Tanaka and M. Okutomi, "Beyond color difference: residual interpolation for color Image demosaicking," *IEEE Transactions on Image Processing*, vol. 25, no. 3, pp. 1288-1300, Mar. 2016.

[15] J. Wu, M. Anisetti, W. Wu, E. Damiani and G. Jeon, "Bayer demosaicking with polynomial interpolation," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5369-5382, Nov. 2016.

[16] Y. Niu, J. Ouyang, W. Zuo and F. Wang, "Low cost edge sensing for high quality demosaicking," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2415-2427, May 2019.

[17] H. S. Malvar, Li-wei He and R. Cutler, "High-quality linear interpolation for demosaicing of Bayer-patterned color images," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Montreal, Canada, May 2004, pp. iii - 485-488.

[18] L. Miao, H. Qi, R. Ramanath and W. E. Snyder, "Binary tree-based generic demosaicking algorithm for multispectral filter arrays," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3550-3558, Nov. 2006.

[19] C. Zhang, Y. Li, J. Wang and P. Hao, "Universal demosaicking of color filter arrays," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5173-5186, Nov. 2016.

[20] D. S. Tan, W. Chen and K. Hua, "DeepDemosaicking: adaptive image demosaicking via multiple deep fully convolutional networks," *IEEE Transactions on Image Processing*, vol. 27, no. 5, pp. 2408-2419, May 2018.

[21] V. Gruev, R. Perkins and T. York, "CCD polarization imaging sensor with aluminum nanowire optical filters," *Optics Express*, vol. 18, no. 18, pp. 19087-19094, Aug. 2010.

[22] S. Mihoubi, P. J. Lapray and L. Bigué, "Survey of demosaicking methods for polarization filter array images," *Sensors*, vol. 18, no. 11, pp. 3688-3707, Nov. 2018.

[23] S. Gao and V. Gruev, "Bilinear and bicubic interpolation methods for division of focal plane polarimeters," *Optics Express*, vol. 19, no. 27, pp. 26161-26173, Dec. 2011.

[24] J. Zhang, H. Luo, B. Hui and Z. Chang, "Image interpolation for division of focal plane polarimeters with intensity correlation," *Optics Express*, vol. 24, no. 18, pp. 20799-20807, Sep. 2016.

[25] S. Gao and V. Gruev, "Gradient-based interpolation method for division-of-focal-plane polarimeters," *Optics Express*, vol. 21, no.1, pp. 1137-1151, Jan. 2013.

[26] P. Thévenaz, T. Blu and M. Unser, "Image interpolation and Resampling" in *Handbook of Medical Imaging*, Washington, USA: SPIE Press, Oct. 2000, pp. 393-420.

[27] B. M. Ratliff, C. F. LaCasse and J. S. Tyo, "Interpolation strategies for reducing IFOV artifacts in microgrid polarimeter imagery," *Optics Express*, vol. 17, no. 11, pp. 9112-9125, Jun. 2009.

[28] S. Gao and V. Gruev, "Gradient based interpolation for division of focal plane polarization imaging sensors," in *Proc. IEEE International Symposium on Circuits and Systems*, Seoul, Korea, May 2012, pp. 1855-1858.

[29] J. Zhang, W. Ye, A. Ahmed, Z. Qiu, Y. Cao and X. Zhao, "A novel smoothness-based interpolation algorithm for division of focal plane Polarimeters," in *Proc. IEEE International Symposium on Circuits and Systems*, Baltimore, MD, May 2017, pp. 1-4.

[30] A. Ahmed, X. Zhao, V. Gruev, J. Zhang and A. Bermak, "Residual interpolation for division of focal plane polarization image sensors," *Optics Express*, vol. 25, no. 9, pp. 10651-10662, May 2017.

[31] J. Zhang, H. Luo, R. Liang, A. Ahmed, X. Zhang, B. Hui and Z. Chang, "Sparse representation-based demosaicing method for microgrid polarimeter imagery," *Optics Letters*, vol. 43, no. 14, pp. 3265-3268, Jul. 2018.

[32] X. Zeng, Y. Luo, X. Zhao and W. Ye, "An end-to-end fully-convolutional neural network for division of focal plane sensors to reconstruct S_0, DoLP, and AoP", *Optics Express*, vol. 27, no. 6, pp. 8566-8577, Mar. 2019.

[33] E. Gilboa, J. P. Cunningham, A. Nehorai and V. Gruev, "Image interpolation and denoising for division of focal plane sensors using Gaussian processes," *Optics Express*, vol. 22, no. 12, pp. 15277-15291, Jun. 2014.

[34] A. Abubakar, X. Zhao, S. Li, M. Takruri, E. Bastaki and A. Bermak, "A block-matching and 3-D filtering algorithm for Gaussian noise in DoFP polarization images," *IEEE Sensors Journal*, vol. 18, no. 18, pp. 7429-7435, Sep. 2018.

[35] K. He, J. Sun and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397-1409, Jun. 2013.

[36] N. Draper and H. Smith, *Applied Regression Analysis*. New York, USA: Wiley, 1981.

[37] C. A. Hall and W. W. Meyer, "Optimal error bounds for cubic spline interpolation," *Journal of Approximation Theory*, vol. 16, no. 2, pp. 105-122, 1976.

[38] N. Li, Y. Zhao, Q. Pan and S. G. Kong, "Demosaicking DoFP images using Newton's polynomial interpolation and polarization difference model," *Optics Express*, vol. 27, no. 2, pp. 1376-1391, Jan. 2019.

[39] Polarization Images [Online]. Available: https://pan.baidu.com/s/1ujcvJ_N6ztWCijBaBQsFkA, code: x96p.

[40] W. Zhou, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image Qualifty Assessment: From Error Visibility to Structural Similarity." *IEEE Transactions on Image Processing*. vol. 13, no. 4, pp. 600-612, Apr. 2004.

[41] A. Hiroshi, "A method of bivariate interpolation and smooth surface fitting based on local procedures." *Communications of the ACM*, vol. 17, no. 1, pp. 18-20, 1974.

[42] H. S. Malvar, Li-wei He and R. Cutler, "High-quality linear interpolation for demosaicing of Bayer-patterned color images," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Montreal, Canada, 2004, pp. 485-488.

[43] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician* , vol. 46, no. 3, pp. 175–185, 1992.

[44] C. N. Ochotorena and Y. Yamashita, "Anisotropic Guided Filtering," *IEEE Transactions on Image Processing*, vol. 29, pp. 1397-1412, Sep. 2019.

[45] P. J. Lapray, L. Gendre, A. Foulonneau and L. Bigué, "Database of polarimetric and multispectral images in the visible and NIR regions," *Unconventional Optical Imaging*, vol. 10677, May 2018.

[46] K. Zhang, W. Zuo, Y. Chen, D. Meng and L. Zhang. "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising." *IEEE Transactions on Image Processing*. vol. 26, no. 7, Feb. 2017, pp. 3142-3155

[47] J. Kim, J. K. Lee and K. M. Lee. "Accurate Image Super-Resolution Using Very Deep Convolutional Networks." In *Proceedings of the IEEE® Conference on Computer Vision and Pattern Recognition*. Nov. 2016, pp. 1646-1654.

Liu Shumin received his PhD in 2019 and B. Eng. with Hons in 2015, both from the Pillar of Engineering Product Development, Singapore University of Technology and Design, Singapore. He is currently a post-doctor in the School of Navigation, Northwestern Polytechnical University, China. His research interests include image processing, multi-focus image fusion, homography estimation for sports video analysis technologies, image denoise, demosaicking and compression image fusion, homography estimation for sports video analysis technologies, image denoise and image compression.

Jiajia Chen received his B. Eng. (Hons) and Ph.D. from Nanyang Technological University, Singapore, in 2004 and 2010, respectively. From April 2012 to March 2018, he was a faculty member in Singapore University of Technology and Design. Since April 2018, he has been with Nanjing University of Aeronautics and Astronautics, China, where he is currently a Professor. His research interest includes computational transformations of low-complexity digital circuits and digital signal processing. Dr. Chen served as Web Chair of Asia-Pacific Computer Systems Architecture Conference 2005, Technical Program Committee member of European Signal Processing Conference 2014 and The Third IEEE International Conference on Multimedia Big Data 2017, and Associate Editor of Springer EURASIP Journal on Embedded Systems since 2016.

Yuan Xun is currently an undergraduate student majoring in Information Engineering at Nanjing University of Aeronautics and Astronautics, Nanjing, China. She has won China's national scholarship in 2018. Her research interest includes computer vision.

Xiaojin Zhao received the B.Sc. degrees in both microelectronics and applied mathematics from Peking University, Beijing, China, in 2005, and the Ph.D. degree in electrical and electronic engineering from The Hong Kong University of Science and Technology (HKUST), Hong Kong, in 2010.

From 2010 to 2011, he was a Postdoctoral Research Associate with HKUST. In 2012, he joined Shenzhen University, Shenzhen, China, where he is currently an Associate Professor with the College of Electronics and Information Engineering. In 2014, he was a Visiting Scholar with IMEC, Leuven, Belgium. He has published over 80 international journal articles and peer-reviewed conference papers. His research interests include CMOS monolithic polarization image sensor, gas sensor, and their related hardware security techniques (e.g., physical unclonable function and true random number generator) when applied to the field of smart Internet of Things (IoT).

Dr. Zhao has served as an organizing/technical committee member in various IEEE conferences. He also serves as a Technical Committee Member for the IEEE Circuits and Systems Society (CASS) on Sensory Systems. He was a co-recipient of the Best Student Paper Award from the IEEE EDSSC'2018 in Shenzhen, China, and the Outstanding Student Paper Award from the IEEE MEMS'2020 in Vancouver, BC, Canada. He has one paper selected as the cover page of the IEEE Electron Device Letters (EDL) in 2020 and two articles selected as the finalist of the Best Paper Award for AsianHOST'2017 in Beijing, and AsianHOST'2019 in Xi'an, China. He has served as the Vice Chair and the Chair for the IEEE Electron Devices and Solid-State Circuits (EDSSC) Shenzhen Joint Chapter from 2015 to 2019.

Chip-Hong Chang (S'92-M'98-SM'03-F'18) received the B.Eng. (Hons.) degree from the National University of Singapore in 1989, and the M. Eng. and Ph.D. degrees from Nanyang Technological University (NTU) of Singapore in 1993 and 1998, respectively. He is an Associate Professor of the School of Electrical and Electronic Engineering (EEE) of NTU. He held joint appointments with the university as Assistant Chair of Alumni from 2008 to 2014, Deputy Director of the Center for High Performance Embedded Systems from 2000 to 2011, and Program Director of the Center for Integrated Circuits and Systems from 2003 to 2009. He has edited 5 books, published 13 book chapters, more than 100 international journal papers (two-thirds are IEEE) and more than 180 refereed international conference papers (mostly in IEEE), and delivered over 40 colloquia. His current research interests include hardware security, unconventional number systems, and low-power and fault-tolerant digital signal processing algorithms and architectures.

Dr. Chang serves as the Associate Editor of IEEE Transactions on Very Large Scale Integration (VLSI) Systems since 2011, IEEE Access from 2013 to 2019, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems and IEEE Transactions on Information Forensics and Security from 2016 to 2019, IEEE Transactions on Circuits and Systems-I since 2020 and from 2010 to 2013, Integration, the VLSI Journal from 2013 to 2015, Springer Journal of Hardware and System Security since June 2016 and Microelectronics Journal since May 2014. He guest edited several IEEE journal special issues and served in the organizing and technical program committee for more than 60 international conferences. He is also an IET Fellow and 2018-2019 Distinguished Lecturer of IEEE Circuits and Systems Society.