西北农林科技大学
NORTHWEST A&F UNIVERSITY

2023 Undergraduate Thesis

Research on DoFP Demosaicking Algorithms

College：	College of Science
Major：	Optoelectronic Information Science and Engineering
Class：	Class of 1902
Student Name：	Yifan Lu
Advisor：	Wenyi Ren
Co-advisor：	None
Completion Date：	Jun 2023

Research on DoFP Demosaicking Algorithms

Abstract： In recent years, polarimetric imaging technology has received widespread attention. Focal plane polarimetric cameras have become a research hotspot due to their compact size, affordability, and ability to acquire polarization information in a single capture. Their distinctive feature lies in the interleaved arrangement of micro-polarizer arrays with varying polarization angles, allowing each pixel to obtain only one polarization angle information after a single capture. Hence, it becomes essential to reconstruct the complete polarization information based on the properties within and between polarization channels. This process is referred to as polarization image demosaicking. This paper comprehensively examines and compares 22 existing demosaicking algorithms based on interpolation theory. These algorithms are categorized into five major groups according to their characteristics, encompassing nearly all interpolation-based polarization image demosaicking approaches. The principles, derivations, and implementation processes of the primary algorithms are thoroughly explained. Finally, using MATLAB and publicly accessible polarized image databases, computer simulation experiments for all algorithms are conducted. Quantitative evaluations using both objective and subjective metrics reveal that residual-based methods and inter-channel correlation interpolation often yield superior results. Notably, the Newton polynomial interpolation and edge-aware residual interpolation methods demonstrate particularly outstanding performance.

Catalog

Chapter 1 Introduction

1.1 Research significance and background

1.1.1 Polarized light

Light carries three main components of information: intensity, frequency, and polarization. However, early color imaging techniques could only capture and utilize the first two types of information. With the advancement of technology, color imaging has become insufficient to meet the increasing demand for light information. As a result, researchers have gradually begun developing imaging systems capable of capturing polarization information. This is due to the wide-ranging applications of polarization information in various aspects of production and daily life. Particularly, parameters such as polarization degree and polarization angle have proven invaluable for understanding surface materials, material structures, 3D reconstruction, image dehazing, and more. Furthermore, polarization imaging can accurately capture information even in scenarios with low light intensity that would otherwise result in unclear imaging. These advantages underscore the immense research value of acquiring and utilizing polarization information in fields such as defense, military, agriculture, industry, and healthcare (Xin et al., 2023).

1.1.2 DoFP demosaicking

Polarization imaging systems analyze the propagation direction and vibration direction of light waves. These systems typically consist of imaging systems combined with polarization optical devices such as polarizers, liquid crystal modulators, and prisms. Recent comprehensive reviews and advancements in polarimetry can be found in related literature (Tyo et al. 2006; Lapray et al. 2018). Existing polarization imaging systems are categorized into time-division polarimetry and simultaneous polarimetry based on whether they capture all polarization information in a single image. Time-division polarimetry requires rotating polarizers to obtain different polarization states of the same scene. However, this method demands consistency in the scene's polarization information during the acquisition time, which is less suitable for capturing dynamic scenes. In terms of imaging system structure, polarization imaging systems are classified into aperture-division, amplitude-division, and division of focal plane (DoFP) types. The DoFP approach, due to its simplicity and ease of integration, has been widely researched (Zhang et al. 2022).

DoFP sensors employ a polarization filter array (PFA) to capture different polarization direction information. The structure is illustrated in Figure 1-1. This design draws inspiration from the Bayer pattern commonly used in earlier color cameras (similar structures using different types of micro-polarizers are employed in multi-channel cameras to balance sampling frequency and different channel information, enabling the uniform acquisition of various channel polarization information in a single shot). A common PFA is composed of superpixels, with each superpixel comprising a 2×2 array of individual pixels. These four pixels are arranged systematically to detect the intensity of light for four different polarization directions (0°, 45°, 90°, and 135°). The arrangement varies among different polarization cameras. The pattern presented in this paper is from the widely used SONY IMX250MZR polarization camera's PFA array, which is used as the basis for subsequent computer simulation demosaicking algorithms.

The remaining polarization information, such as polarization angle and polarization degree, can be calculated from the intensity values of these four polarization directions. Not every pixel can acquire the complete polarization information in a single shot. One approach is to combine the polarization information of these four pixels (labeled as A, B, C, and D) and treat them as the polarization information of the central point (labeled as F) in a 2×2 region, and then calculate the polarization angle and degree of point F. However, this method introduces significant instantaneous field of view (IFoV) errors and compromises resolution due to the spatial disparity between point F and points A, B, C, and D. To address this, each pixel's other three polarization directions need to be reconstructed from its original location on the focal plane. This process is known as polarization demosaicking. The resulting image after demosaicking contains artifacts such as zipper noise and false edges compared to the ground truth, thus driving the research for fast and accurate demosaicking algorithms.

In the following, we briefly summarize the domestic and international research on demosaicking algorithms.
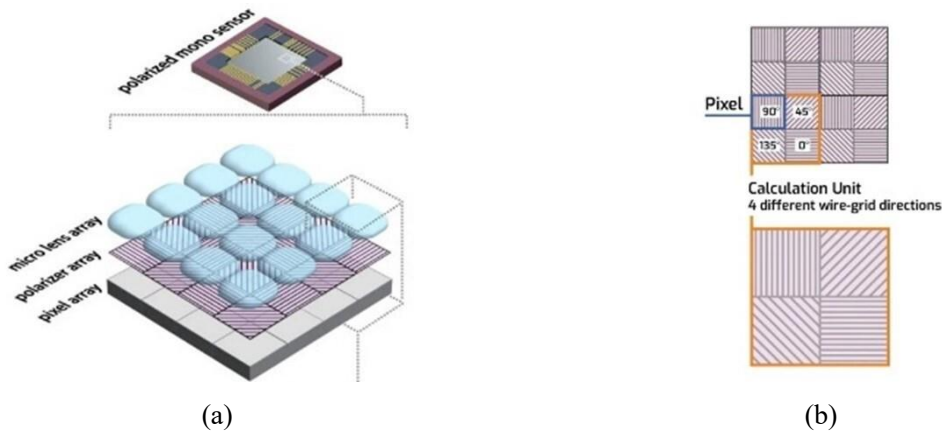
(a)   (b)

Image source: Sony LUCID camera official website（https://thinklucid.com/tech-briefs/polarization-explained-sony-polarized-sensor/）

Fig. 1-1 PFA pattern and structure

1.2 Polarization Demosaicking: Domestic and International Perspectives

Earliest color cameras utilized color filter arrays (CFAs), with the Bayer pattern being the most widely used. Numerous demosaicking algorithms have been developed for CFAs; however, due to differences in arrangement between CFA arrays and polarization filter arrays (PFAs), demosaicking algorithms for CFAs cannot be directly applied to polarization images.

In recent years, various algorithms aimed at improving the image quality, reducing instantaneous field of view (IFoV) errors, image artifacts, and zipper noise, while enhancing the peak signal-to-noise ratio (PSNR), have been proposed for polarization image demosaicking. Interpolation methods are one category of demosaicking algorithms that are intuitive and fast. They interpolate unknown pixel values based on known points in the vicinity. The general approach involves selecting points for interpolation, calculating corresponding weights, and performing weighted averaging to solve for the unknown pixel information.

Traditional interpolation-based polarization image demosaicking algorithms include nearest-neighbor, bilinear, bicubic, and cubic-spline interpolation. These basic interpolation methods rely on convolving and upsampling the original image using linear low-pass filters, which, however, can result in the loss of high-frequency components (edges, details, etc.) in the image.

Many subsequent algorithms build upon these basic methods, utilizing inter-channel and intra-channel correlations in polarization images. They employ different approaches to calculate weights for known points, aiming to enhance interpolation accuracy for high-

frequency components, particularly edges, and thus improve both subjective and objective evaluation metrics.

Tyo et al. (2009) improved upon traditional bilinear interpolation by proposing four enhanced convolution kernels that consider the distance between points as a weighting factor. Building on this, an adaptive convolution kernel was designed by incorporating intensity weighting (2011). Gao and Gruev (2013) separated interpolation along edges from that in smooth regions by utilizing image gradients for edge detection. Zhang et al. (2016) also separated interpolation for edges and smooth regions, combining inter-channel correlations to determine edges with richer information than previous methods. Kiku et al. (2013) introduced an algorithm applicable to RGB image demosaicking, and its concept of residual channel interpolation can be transposed to polarization image algorithms. Since the residual channel contains fewer high-frequency components, this approach leads to lower interpolation errors in that channel. Li et al. (2019) applied inter-channel interpolation, combining Newton polynomial interpolation and Taylor expansion for error estimation, and gradient calculation for edge detection, achieving excellent demosaicking results. Morimatsu et al. (2020) employed different gradient definitions and judgements, along with guided filtering, to achieve accurate reproduction at low complexity. Mihoubi et al. (2015) defined an image considering all four channels, resulting in smoother residual channels and improved subsequent interpolation accuracy. Liu et al. (2020) employed an edge compensation approach for pre-interpolation, combining interpolation and inter-channel correlations.

Furthermore, some frequency-domain transformation-based interpolation methods exist, such as wavelet transform (Wang et al. 2015). These methods also rely on the strong correlation of high-frequency components between channels for interpolation. The frequency-domain advantage lies in the direct extraction and replacement of high-frequency information, enabling high-frequency restoration across all channels, followed by low-frequency restoration.

In recent years, many demosaicking algorithms incorporating deep learning have surpassed traditional methods. However, deep learning algorithms often require suitable databases, extensive training, data calibration, and long computation times. Additionally, most commercially available polarization imaging hardware does not fully support such algorithms. As a result, this paper does not delve into this approach.

1.3 Main Content and Organization of This Thesis

This thesis primarily investigates traditional interpolation algorithms for demosaicking based on polarized focal plane images. It covers most of the typical algorithms that have been extensively researched since the study of polarized demosaicking algorithms began, including the innovative algorithms mentioned in Section 1.2. Each algorithm is comprehensively introduced in this paper, and they are implemented using MATLAB programs written by the author or provided by the original authors. The algorithms are tested using a database, and representative objective and subjective evaluation metrics are selected for assessment, discussing their advantages and limitations.

Chapter 1 introduces the background and significance of the selected research topic, as well as the domestic and foreign research status of interpolation-based polarized demosaicking algorithms.

Chapter 2 describes the mathematical representation of the polarization properties of light, introduces the structure and imaging characteristics of DoFP polarimeters, leading to the basic process of using interpolation for demosaicking. Finally, it introduces possible errors and image quality issues in mosaic restoration, as well as common evaluation metrics for demosaicked images.

Chapter 3 studies classical interpolation-based demosaicking algorithms. It first introduces important properties of polarized images, combined with statistical results from a large number of polarized images in the database to illustrate the strong correlation between channels and within channels. This serves as the fundamental principle for interpolation-based demosaicking. It then introduces 22 typical interpolation algorithms and their implementation steps from 2006 to the present, categorizing them into five major groups based on their characteristics.

Chapter 4 demonstrates the demosaicking performance of the algorithms. Computer simulations are conducted on all the aforementioned algorithms using publicly available databases, and the most commonly used objective polarization evaluation metrics are used as statistical standards for comparison. The algorithms are analyzed in comparison with conclusions from other review literature and subjective evaluation metrics.

Chapter 5 summarizes the work and content of this thesis, and provides an outlook for future research.

Chapter 2 Basic Theory

Before presenting the polarized demosaicking algorithms within the system, it is necessary to supplement some fundamental theories. We need to understand how individuals quantify polarization information of light using methods such as matrices, and how they utilize this aspect to design specific hardware structures for information acquisition. This chapter introduces the characteristics of polarization imaging systems under different design approaches. This, in turn, provides an intuitive introduction to the rationale and process behind the existence of demosaicking algorithms. As a means of image restoration, algorithms of this kind inherently involve a comparison between the restored and original images. Hence, the final part of this chapter introduces commonly used image evaluation metrics.

2.1 Matrix representation of polarized light

2.1.1 Jones Matrix

Jones introduced the concept of Jones matrices in 1941, providing a mathematical description of polarization light amplitudes and phases. In this method, a beam of light is represented as a vector containing two complex elements, and a linear polarizing system is represented by a 2×2 matrix known as the Jones matrix. If a polarized light A passes through a polarizing element B to yield polarized light C, it can be expressed through matrix multiplication as $AB = C$. It's important to note that Jones matrices can only represent completely polarized light and cannot describe natural light or partially polarized light.

The specific representation of Jones matrices is as follows:

$$E = E_x \cdot X_0 + E_y \cdot Y_0 = E_x \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + E_y \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} E_x e^{i\delta_x} \\ E_y e^{i\delta_y} \end{bmatrix} \qquad (2-1)$$

where E represents the optical field and is a vector quantity. The commonly used Jones matrix is the normalized：

$$E = \begin{bmatrix} \cos\alpha \\ e^{i\delta}\sin\alpha \end{bmatrix} \qquad (2-2)$$

where $\cos\alpha = \frac{E_x}{\sqrt{E_x^2+E_y^2}}$, $\sin\alpha = \frac{E_y}{\sqrt{E_x^2+E_y^2}}$, are the normalized light projections on the X, Y axis respectively, and $\alpha \in \left(0, \frac{\pi}{2}\right)$.

2.1.2 Stokes Matrix

Stokes proposed the Stokes vector in 1852 to describe polarized light.

The matrix can describe not only fully polarized light, but also incoherent light, partially polarized light, which is a method that can describe polarized light more comprehensively. It can represent the polarization system with a Mueller matrix, which is a generalization of Jones matrix, as a 4×4 matrix.

The specific representation of the Stokes matrix is as follows (Shaw 2016)：

$$S = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} \langle |E_x|^2 + |E_y|^2 \rangle \\ \langle |E_x|^2 - |E_y|^2 \rangle \\ 2\mathrm{Re}\langle E_x E_y^* \rangle \\ -2\mathrm{Im}\langle E_x E_y^* \rangle \end{bmatrix} \qquad (2-3)$$

where E_x, E_y represent the light field components, S_0 represents the total light intensity, S_1 is the difference between the horizontal and vertical components of polarized light, S_2 is the difference between the components of polarized light vibrating along 45° and 135° respectively, S_3 is the difference between the right- and left-handed components of light, and since this paper only considers linearly polarized light, there is no phase difference between the components, so S_3 is always 0.

2.1.3 Physical quantities describing polarization characteristics

As this paper focuses on fully polarized light excluding elliptical or circular polarization, where all components in the light field have the same phase, the following will introduce the physical quantities used to describe linearly polarized light. Moreover, all of these quantities can be represented using components of the Stokes vector.

Angle of Linear Polarization (AoLP), To describe linearly polarized light on the plane of vibration, the angle φ between the direction of vibration and the horizontal axis is given by the following mathematical expression

$$\varphi = 0.5 \arctan \frac{S_2}{S_1} \qquad (2-4)$$

Degree of Linear Polarization (DoLP), which represents the proportion of linearly polarized light in the total light intensity, has the following mathematical expression

$$DoLP = \sqrt{S_1^2 + S_2^2}/S_0 \qquad (2-5)$$

2.2 Polarized light imaging system

2.2.1 Polarization imaging sensor classification

There are several types of polarization detection devices available on the market, primarily categorized into four major groups: time-division polarimeters, amplitude-division polarimeters, aperture-division polarimeters, and division of focal plane (DoFP) polarimeters (Shaw, 2016).

The structure and usage of time-division polarimeters are intuitive, and they come at a lower cost. However, maintaining relative stability between the scene and the platform is necessary to avoid misalignment between frames. This type of polarimeter modulates the polarized light emitted by the scene on the focal plane by rotating polarization elements. The modulated data can be used to reconstruct polarization information.

On the other hand, amplitude-division polarimeters are dual-channel instruments capable of detecting all Stokes parameters. They consist of four separate focal planes, a polarizing beam splitter, a phase retarder, lenses, and other components. This system has a larger volume and requires high mechanical precision during assembly. Due to the high degrees of freedom during installation and adjustment, deviations between polarization channels can occur, necessitating subsequent registration of the four images.

Aperture-division polarimeters can simultaneously acquire all polarization images while ensuring consistent fields of view for all channels. They use a single focal plane and an imaging system to calibrate multiple images onto the same focal plane, eliminating the need for subsequent image alignment. However, this approach may sacrifice spatial resolution and involve larger volume and weight, making precise assembly challenging.

The division of focal plane (DoFP) polarimeters are less sensitive to phase differences. However, designs for full Stokes matrix detection in DoFP systems have been proposed (Nordin et al., 1999). DoFP systems have been applied in the full spectral range. Most DoFP systems feature an interleaved arrangement of polarizers, as shown in Figure 1-1. These systems achieve the simultaneous acquisition of different polarization angle channel information by reducing the spatial sampling frequency. They reconstruct the Stokes matrix at the center position by utilizing the polarized light information from four neighboring directions. However, this approximation introduces instantaneous field of view (IFoV) errors and pixel misalignment, as these adjacent points do not perfectly coincide spatially. The introduction of demosaicking

algorithms can mitigate IFoV errors and minimize the loss of polarization information.

2.2.2 DoFP-based imaging analysis

Based on the above analysis, the DoFP (Degree of Polarization) system acquires more information simultaneously by balancing spatial sampling frequency and information from different polarization angle channels. To achieve the best balance and acquisition efficiency, the design of the focal plane as a polarization array needs to be considered.

To achieve a higher signal-to-noise ratio, selecting the optimal polarization angles for measurements should be uniformly distributed between 0 and 180 degrees (Zhang et al., 2022). The DoFP polarization array applies sampling angles of 0°, 45°, 90°, and 135° as vibration directions. Without considering left and right circular polarization, the method for solving the Stokes matrix of a light beam can be transformed into

$$\begin{cases} S_0 = 0.5(I_0 + I_{45} + I_{90} + I_{135}) \\ S_1 = I_0 - I_{90} \\ S_2 = I_{45} - I_{135} \end{cases} \qquad (2-6)$$

where I_0, I_{45}, I_{90} and I_{135} indicate the light intensity obtained after the beam passes through 0°, 45°, 90° and 135° miniature polarizers, respectively.

Due to the structure of the DoFP (Degree of Polarization) micro-polarization array, in order to obtain the polarization information of the entire planar beam while avoiding the aforementioned IFoV (Instantaneous Field of View) misalignment phenomenon, the solution is to acquire the I_0, I_{45}, I_{90} and I_{135} information for each pixel. Based on the structure of the DoFP array, it is known that directly obtaining a single polarization angle information is possible through a single pixel on the polarization plane. Therefore, it is necessary to use an algorithm to reconstruct the other three polarization intensities at each point in order to obtain the polarization information of the entire planar pixel. This reconstruction algorithm is known as the focal plane polarization demosaicking algorithm, and the quality of the algorithm directly affects the accuracy of the subsequently computed polarization parameters. Therefore, proposing a robust demosaicking algorithm is of paramount importance.

2.3 Demosaicking flow based on DoFP

The earlier mentioned DoFP focal plane is composed of interleaved arrays of polarization elements, and demosaicking requires the reconstruction of the other three polarization information values for each pixel using surrounding points. There are various methods to

achieve this process, and this paper primarily focuses on the interpolation method. This method is the most intuitive, classical, efficient, and aligned with the current mainstream hardware for polarization imaging demosaicking. The general operational process is as follows (Mihoubi et al. 2018):

The PFA camera acquires an image I_{raw} of size $X \times Y$ from the focal plane, where each pixel p will contain a vibration direction k (0°, 45°, 90° or 135°). Assuming that S is the full set containing all directional polarization information of all pixel points, i.e., a matrix of $X \times Y \times 4$, and S^k is the set of information of which the vibration direction is k, i.e., a subset of S, a matrix of X×Y, then $S = \bigcup_{k \in \{0,45,90,135\}} S^k$.

I_{raw} can be viewed as the result of combining different S^k after downsampling at pixel, i.e., $\forall p \in S, I_p^{raw} = I_p^{PFA(p)}$. I_{raw} can also be viewed as the sum of the sparse matrices of the four channels, which is the $X \times Y$ matrix obtained after down-sampling each S^k, which includes the sampled point values and the non-sampled point positions values (for 0): $I^k = I_{raw} \odot m_k$, where

$$m_p^k = \begin{cases} 1 & if\ PFA(p) = k, i.e., p \in S^k \\ 0 & otherwise \end{cases}$$

。

Demosaicking is the process of filling in pixel values at non-sampled positions on each sparse matrix, thereby obtaining complete 4-channel values. For each pixel point p, the values of the other three channels are estimated through interpolation. The demosaicking process is illustrated in Figure 2-1.
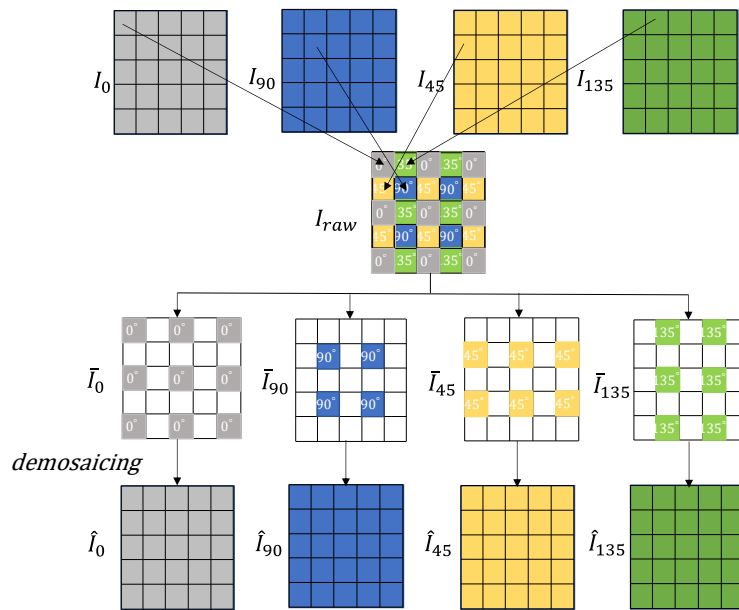

Fig. 2-1 Flowchart of the demosaicking algorithm

2.4 Image Evaluation Metrics

Image evaluation metrics include subjective and objective evaluation metrics, and since four groups of images exist for comparison before and after demosaicking (0°, 45°, 90°, and 135°), we discuss a channel separately, and the other channels are similar. It is assumed that there is a difference between the demosaicked image D and the real image I. Subjective evaluation involves human observation of the differences between two images, particularly focusing on edges and fine details. It assesses the degree of artifacts such as false edges and zipper noise, qualitatively analyzing the algorithm's results. Objective metrics quantitatively measure the magnitude of differences between the two images, enabling a comparison of algorithm performance.

We use the Peak Signal-to-noise Ratio (PSNR), a commonly used image evaluation metric, as the subsequent objective evaluation metric in this paper, and the PSNR expression between two images of the same size (I and D) is as follows (Mihoubi et al. 2018)

$$PSNR(I, D) = 10\log_{10}\left(\frac{R^2}{MSE(I, D)}\right) \qquad (2-7)$$

where R is the maximum fluctuation value of the image I's data type, which is 255 in the case of an 8 bit image depth. MSE is the mean squared error between image I and D, which is expressed as follows

$$MSE = \frac{1}{mn} \sum_{1 \leq i \leq m} \sum_{1 \leq j \leq n} \left(I(i,j) - D(i,j)\right)^2 \qquad (2-8)$$

where m and n are the number of rows and columns of the image respectively.

It is easy to see that MSE reflects the absolute value of the overall average error of the recovered image, and the smaller it is, the better the image quality; PSNR reflects the relative value of the inverse of the maximum error of the recovered image, and the larger it is, the better the image quality is, and the effect of the algorithm can be compared under different databases, which is more widely used.

2.5 Summary of this chapter

This chapter primarily introduces several commonly used matrix representations of polarized light, such as Jones matrix and Stokes matrix, as well as prevalent polarimetric parameters like linear polarization angle and linear polarization degree. The chapter discusses

the advantages and disadvantages of polarization imaging systems under different design approaches, emphasizing the strengths of focal plane polarization imaging systems. Consequently, it derives the necessity and process of polarization image demosaicking algorithms based on their characteristics. Finally, the chapter introduces the widely-used image evaluation metric, PSNR, as an objective measure to assess the restoration effectiveness of subsequent algorithms.

Chapter 3 Comparison of the principles in traditional polarization image demosaicking algorithms

In the preceding content, the importance of demosaicking algorithms was highlighted based on the characteristics of polarized light and polarized imaging hardware. The concept of polarized image demosaicking through interpolation is not arbitrarily conceived; rather, it follows a logical continuity from the principles of demosaicking for color images. The underlying principles are similar. Regardless of the image type, the design and innovation of demosaicking algorithms are founded on the inter-channel and intra-channel correlations within images. This chapter begins by presenting statistical research on image correlations. Building upon this foundation, the chapter systematically elaborates on classic interpolation methods applied to polarized image demosaicking, tracing their development from underlying principles, logical inspiration, and innovative logic.

3.1 Study of channel correlation

Whether it is CFA, SFA or PFA, the interpolation method unfolds with inter- and intra-channel phase correlation of the image channels.

Intra-channel correlation: The interpolation algorithm considers that the pixel points belonging to homogeneous regions are strongly correlated with their neighboring points. Therefore, a picture can be considered as a combination of scattered homogeneous regions split by edges, and all points inside the homogeneous regions can be better estimated using their neighborhoods. For edges, most algorithms use the gradient to make a judgment, and the larger the gradient, the more likely it is that the vertical direction is an edge, and the less likely it is that these two neighboring points are within a homogeneous region (Mihoubi et al. 2018).

Inter-channel correlation: the interpolation method considers that the images under different channels of a scene are highly correlated with their high-frequency part information (edges, details, etc.), so the channel with higher spatial sampling frequency (if it exists) should be recovered as a priority, which ensures that the high-frequency information is reproduced as well as possible, and then the high-frequency information of this recovered channel is used to guide the other channels to recover their high-frequency parts (Mihoubi et al . 2018).

Experimental validation: in their literature, Mihoubi et al. (2018) used polarization images from a database (Lapray et al. 2018) to perform statistics and to validate the two properties mentioned above.

They performed two sets of controlled variable experiments. The first group set different polarization angles 0°, 45°, 90° and 135° and observed the inter-channel correlation, intra-channel correlation (spatial correlation) of six different wavelength channels (five visible and one near-infrared light): it was found that regardless of the polarization angle, excluding the channel with high noise influence, there was always a strong correlation between the six channels, and the mean value of intra-channel spatial correlation was as high as 0.9504. Meanwhile the closer the wavelengths or polarization angles, the stronger the correlation between the channels. The second group uses wavelength as the variable to observe the inter-channel and intra-channel correlations of channels with different polarization angles, and concludes that the inter-channel and intra-channel correlations of polarization angles are always higher regardless of the wavelength; and the closer the polarization angle is, the stronger the correlation of the channels.

The conclusion of this experiment verifies the theoretical basis mentioned before, and also lays the foundation for the subsequent interpolation of the polarization demosaicking algorithm.

3.2 DoFP-based interpolation algorithm

3.2.1 Spatially invariant linear low-pass filter

（1）Nearest

Nearest neighbor interpolation is the simplest traditional interpolation algorithm, also known as zero-order interpolation. For each pixel, a 2×2 superpixel containing the pixel is framed in a certain direction (lower right direction), taking Figure 3-1 image element A(1,1) as an example, the other three pixel values in the superpixel are treated as the three channel values of A. The expression is as follows (Wan Can 2022):

$$\begin{cases} I_0(1,1) = I_A \\ I_{45}(1,1) = I_B \\ I_{135}(1,1) = I_C \\ I_{90}(1,1) = I_D \end{cases} \qquad (3-1)$$

0°A	45°B	0°	45°	
135° C	90°D	135°	90°	...

Fig. 3-1 Nearest interpolation pixel location diagram

The above demosaicking process can also be expressed in terms of convolution (Ratliff et al. 2009), with the convolution kernel

$$H = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Suppose the single-channel sparse matrix is i_k and this complete channel is I_k after interpolation, then

$$I_k = i_k * H \tag{3-2}$$

This convolution is performed on each of the four channel sparse matrices to obtain the complete demosaiced four-channel image. Since this method simply copies the neighborhood information, it leads to a strong jagged effect.

（2）Bilinear

The bilinear interpolation method is very widely used and is also commonly used as a pre-processing algorithm for other interpolation methods. Its main principle is shown in Figure 3-2, using the values of the four adjacent corner positions of the point to be interpolated (x, y) for weighted summation, the weights are related to their distances from the point (x, y). The process can be viewed as doing two linear interpolations in the x-direction and then one linear interpolation in the y-direction respectively (Wan Can 2022).
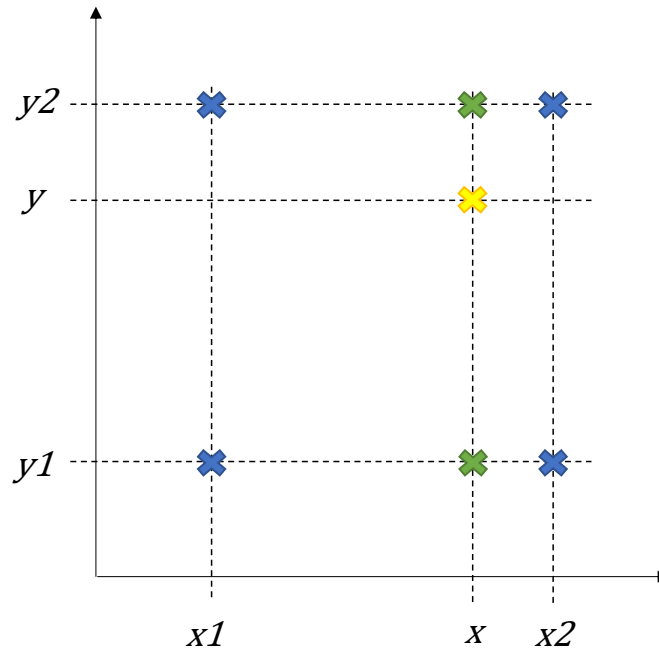

Fig. 3-2 Schematic diagram of pixel location by bilinear interpolation

$$I_{xy} = \frac{(y_2 - y)}{(y_2 - y_1)}\left[\frac{(x_2 - x)}{(x_2 - x_1)}I_{11} + \frac{(x - x_1)}{(x_2 - x_1)}I_{21}\right] +$$
$$\frac{(y - y_1)}{(y_2 - y_1)}\left[\frac{(x_2 - x)}{(x_2 - x_1)}I_{12} + \frac{(x - x_1)}{(x_2 - x_1)}I_{22}\right] \tag{3-3}$$

where I_{11}, I_{12}, I_{21} and I_{22} are the grayscale values of the points (x_1, y_1), (x_1, y_2), (x_2, y_1) and (x_2, y_2), respectively

The above demosaicking process can also be represented by convolution (Ratliff et al. 2009). The convolution kernel is

$$H = 0.25 \times \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Suppose the single-channel sparse matrix is i_k and this complete channel is I_k after interpolation, then

$$I_k = i_k * H \tag{3-4}$$

This convolution is done for all four channels of the sparse matrix to obtain the complete demosaiced four-channel image.

（3）Bicubic

The dual cubic interpolation method is to do three interpolations in each of the two dimensions and apply them to a total of 16 points in the neighboring 4×4 for the operation. The values of the points to be interpolated are derived by constructing functions in the region. Let the known point value be f, and its derivatives be f_x, f_y and f_{xy}, and the four adjacent points be $(0,1)$, $(1,0)$, $(0,0)$ and $(1,1)$, the interpolation function p (a surface passing through these four points) is denoted as

$$p(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij}x^i y^i \tag{3-5}$$

The interpolation problem is transformed into solving the function p with 16 coefficients to be determined. It is also known that:

$$f(0,0) = p(0,0) = a_{00} \tag{3-6}$$

$$f(1,0) = p(1,0) = a_{00} + a_{10} + a_{20} + a_{30} \tag{3-7}$$

$$f(0,1) = p(0,1) = a_{00} + a_{01} + a_{02} + a_{03} \tag{3-8}$$

$$f(1,1) = p(1,1) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} \tag{3-9}$$

$$f_x(0,0) = p_x(0,0) = a_{10} \tag{3-10}$$

$$f_x(1,0) = p_x(1,0) = a_{10} + 2a_{20} + 3a_{30} \tag{3-11}$$

$$f_x(0,1) = p_x(0,1) = a_{10} + a_{11} + a_{12} + a_{13} \tag{3-12}$$

$$f_x(1,1) = p_x(1,1) = \sum_{i=1}^{3}\sum_{j=0}^{3} a_{ij}\, i \tag{3-13}$$

$$f_y(0,0) = p_y(0,0) = a_{01} \tag{3-14}$$

$$f_y(1,0) = p_y(1,0) = a_{01} + a_{11} + a_{21} + a_{31} \tag{3-15}$$

$$f_y(0,1) = p_y(0,1) = a_{01} + 2a_{02} + 3a_{03} \tag{3-16}$$

$$f_y(1,1) = p_y(1,1) = \sum_{i=0}^{3}\sum_{j=1}^{3} a_{ij}\, j \tag{3-17}$$

$$f_{xy}(0,0) = p_{xy}(0,0) = a_{11} \tag{3-18}$$

$$f_{xy}(1,0) = p_{xy}(1,0) = a_{11} + 2a_{21} + 3a_{31} \tag{3-19}$$

$$f_{xy}(0,1) = p_{xy}(0,1) = a_{11} + 2a_{12} + 3a_{13} \tag{3-20}$$

$$f_{xy}(1,1) = p_{xy}(1,1) = \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}\, ij \tag{3-21}$$

Suppose the matrix of unknowns is

$$\alpha = [a_{00}\ a_{10}\ a_{20}\ a_{30}\ a_{01}\ a_{11}\ a_{21}\ a_{31}\ a_{02}\ a_{12}\ a_{22}\ a_{32}\ a_{03}\ a_{13}\ a_{23}\ a_{33}]^{T} \tag{3-22}$$

$$x = [f(0,0)\ f(1,0)\ f(0,1)\ f(1,1)\ f_x(0,0)\ f_x(1,0)\ f_x(0,1)\ f_x(1,1)$$
$$f_y(0,0)\ f_y(1,0)\ f_y(0,1)\ f_y(1,1)\ f_{xy}(0,0)\ f_{xy}(1,0)\ f_{xy}(0,1)\ f_{xy}(1,1)]^{T} \tag{3-23}$$

Then the above equation can be expressed as follows:

$$A^{-1}x = \alpha \tag{3-24}$$

where

$$
A^{-1} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-3 & 3 & 0 & 0 & -2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & -2 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -3 & 3 & 0 & 0 & -2 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 & 1 & 1 & 0 & 0 \\
-3 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -3 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & -1 & 0 \\
9 & -9 & -9 & 9 & 6 & 3 & -6 & -3 & 6 & -6 & 3 & -3 & 4 & 2 & 2 & 1 \\
-6 & 6 & 6 & -6 & -3 & -3 & 3 & 3 & -4 & 4 & -2 & 2 & -2 & -2 & -1 & -1 \\
2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
-6 & 6 & 6 & -6 & -4 & -2 & 4 & 2 & -3 & 3 & -3 & 3 & -2 & -1 & -2 & -1 \\
4 & -4 & -4 & 4 & 2 & 2 & -2 & -2 & 2 & -2 & 2 & -2 & 1 & 1 & 1 & 1
\end{bmatrix} \qquad (3-25)
$$

That is, the unknown coefficients can be derived and thus brought into the p-expression as well as the unknown point coordinates for interpolation (Wikipedia 2023).

Three-time convolution interpolation is an approximate simplified expression of three-time difference (Gao and Gruev 2013), which utilizes segmented function approximations instead of three-time difference functions and uses three-time functions on each segment (Keys 1981):

$$
W(x) = \begin{cases}
(a+2)|x|^3 - (a+3)|x|^2 + 1, & |x| \leq 1 \\
a|x|^3 - 5|x|^2 + 8a|x| - 4a, & 1 < |x| < 2 \\
0, & otherwise
\end{cases} \qquad (3-26)
$$

The common values of a are -0.5 (to ensure that the function W is equal to at least the first three orders of derivatives of the true function), -1 and -0.75, and the parameters are determined according to different needs. Apply it to the two-dimensional space for interpolation, the process is as follows: let the point to be interpolated be at the center of the 4×4 plane, as shown in Figure 3-3 yellow dot, to obtain its 90° information f, the f expression is

$$
f(u,v) = W_u \cdot F \cdot W_v \qquad (3-27)
$$

$$
W_u = [W(u-i+1), W(u-i), W(i-u+1), W(i-u+2)] \qquad (3-28)
$$

$$
F = \begin{bmatrix}
f(i-1,j-1) & f(i-1,j) & f(i-1,j+1) & f(i-1,j+2) \\
f(i,j-1) & f(i,j) & f(i,j+1) & f(i,j+2) \\
f(i+1,j-1) & f(i+1,j) & f(i+1,j+1) & f(i+1,j+2) \\
f(i+2,j-1) & f(i+2,j) & f(i+2,j+1) & f(i+2,j+2)
\end{bmatrix} \qquad (3-29)
$$

$$
W_v = [W(v-j+1), W(v-j), W(j-v+1), W(j-v+2)]^T \qquad (3-30)
$$

90°	135°	90°	135°	90°	135°	90°
45°	0°	45°	0°	45°	0°	45°
90°	135°	90°	135°	90°	135°	90°
45°	0°	45°	0°	45°	0°	45°
90°	135°	90°	135°	90°	135°	90°
45°	0°	45°	0°	45°	0°	45°
90°	135°	90°	135°	90°	135°	90°

Fig. 3-3 Schematic diagram of pixel location by bicubic interpolation

If the desired information is the 45° polarized information at a certain position, the procedure involves performing cubic spline interpolation along the 45° direction for each row. This yields three values. Then, linear interpolation is carried out between adjacent pairs of these three values, resulting in two values. Finally, a linear interpolation between these two values yields the interpolated 45° information at the central position. The same procedure applies to obtaining the interpolated information at 135° (Wan Chan 2022).

（4）Bicubic spline interpolation

Bicubic spline (BS), which also constructs a cubic polynomial at four adjacent points of the unknown point (Gao and Gruev 2011):

$$f_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3 \qquad (3-31)$$

where f_i (between the points x_i and x_{i+1}) is the point to be interpolated, and the following set of equations about the function coefficients can be obtained based on the f_i values, first-order derivatives, and second-order derivatives in succession between the two points:

$$a_i = I(i, j) \qquad (3-32)$$

$$b_i = I(i + 1, j) - I(i, j) - \frac{M_{i+1} - M_i}{6} \qquad (3-33)$$

$$c_i = M_i/2 \qquad (3-34)$$

$$d_i = (M_{i+1} - M_i)/6 \qquad (3-35)$$

$$M_i - 4 * M_{i+1} + M_{i+2} = 6\big(I(i,j) - 2I(i+1,j) + I(i,j)\big) \qquad (3-36)$$

where M_i is the second order derivative of the f_i function

Together with the boundary condition $M_1 = M_n = 0$ the coefficients to be determined for the function can be found. The system of equations is expressed in matrix as follows:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & 4 & 1 & \vdots\vdots\vdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \vdots\vdots\vdots & 1 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 0 & & 0 & 0 & 0 & 1 \end{bmatrix}_{n\times n} \cdot \begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \\ \vdots \\ M_{n-3} \\ M_{n-2} \\ M_{n-1} \\ M_n \end{bmatrix} = \frac{3}{2} \begin{bmatrix} 0 \\ y_1 - 2y_2 + y_3 \\ y_2 - 2y_3 + y_4 \\ y_3 - 2y_4 + y_5 \\ \vdots \\ y_{n-4} - 2y_{n-3} + y_{n-2} \\ y_{n-3} - 2y_{n-2} + y_{n-1} \\ y_{n-2} - 2y_{n-1} + y_n \\ 0 \end{bmatrix} \qquad (3-37)$$

$$A \cdot M = Y \xrightarrow{M_0 = M_n = 0} A' \cdot M' = Y' \qquad (3-38)$$

where A is the $(n-2) \times n$ matrix left after removing the top and bottom rows of the $n \times n$ coefficient matrix, A' is the $(n-2) \times (n-2)$ matrix left after removing the left and right columns of matrix A, and M' and Y' are the $(n-2) \times 1$ column matrices left after removing the first and last rows of each of M and Y, respectively.

Since the previous coefficient matrix is a diagonal matrix, the equations of this matrix can be solved easily by Crout or Doolittle method, making the operation faster. The disadvantage of this method is the large memory consumption, which can limit the real-time execution of the algorithm, so the literature (Gao and Gruev 2011) proposes to restrict the interpolation to be performed in a 10×10 window, while the boundary condition becomes $M_0 = M_{10} = 0$, which ensures the independent and simultaneous execution of operations in multiple windows.

（5）Bilinear with 5 kernels

Bilinear with 5 kernels (B5K) was first proposed to apply bilinear interpolation to DoFP. Three bilinear kernels and two weighted bilinear kernels were proposed to reduce the IFoV error generated by DoFP when acquiring information (Ratliff at al. 2009). The patterns and convolution kernels of these five algorithms are shown in Figures 3-4：

method	pattern	kernels				weights
B0		$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$	

B1		$\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix}$	$\dfrac{1}{2}\begin{bmatrix}0&0&0\\1&0&1\\0&0&0\end{bmatrix}$	$\dfrac{1}{2}\begin{bmatrix}0&1&0\\0&0&0\\0&1&0\end{bmatrix}$		
B2		$\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix}$	$\dfrac{1}{2}\begin{bmatrix}0&1&0\\0&0&0\\0&1&0\end{bmatrix}$	$\dfrac{1}{2}\begin{bmatrix}0&0&0\\1&0&1\\0&0&0\end{bmatrix}$	$\dfrac{1}{4}\begin{bmatrix}1&0&1\\0&0&0\\1&0&1\end{bmatrix}$	
B3		$\begin{bmatrix}0&B&0&0\\0&0&0&0\\0&A&0&B\\0&0&0&0\end{bmatrix}$	$\begin{bmatrix}0&0&B&0\\0&0&0&0\\B&0&A&0\\0&0&0&0\end{bmatrix}$	$\begin{bmatrix}0&0&0&0\\B&0&A&0\\0&0&0&0\\0&0&B&0\end{bmatrix}$	$\begin{bmatrix}0&0&0&0\\0&A&0&B\\0&0&0&0\\0&B&0&0\end{bmatrix}$	A=0.4086 B=0.2957
B4		$\begin{bmatrix}0&B&0&C\\0&0&0&0\\0&A&0&B\\0&0&0&0\end{bmatrix}$	$\begin{bmatrix}C&0&B&0\\0&0&0&0\\B&0&A&0\\0&0&0&0\end{bmatrix}$	$\begin{bmatrix}0&0&0&0\\B&0&A&0\\0&0&0&0\\C&0&B&0\end{bmatrix}$	$\begin{bmatrix}0&0&0&0\\0&A&0&B\\0&0&0&0\\0&B&0&C\end{bmatrix}$	A=0.3541 B=0.2639 C=0.1180

Fig. 3-4 Convolutional kernel with 5 types of bilinear interpolation

The yellow dot represents the position of the point to be recovered. The weighted value is determined based on the distance between the point and the unknown point. Using $i = 0, \dots, 4$ to represent the method serial numbers and $j = 1, \dots, 4$ to represent the serial numbers of the nuclei under each method in order from left to right, the recovery process for each channel is as follows:

$$I_{ij} = P * H_{ij} \tag{3-39}$$

where P represents the directly acquired mosaic image, H_{ij} represents the jth kernel of method i, and I_{ij} represents the jth channel reproduced under that method. It is worth noting that method 1 has only three kernels, i.e., only three channels can be reproduced by the above equation, and the fourth channel is derived using the redundancy of the information acquired by DoFP (Ratliff at al. 2007).

$$\bar{I}_0 = I_{45} - I_{90} + I_{135} \tag{3-40}$$

$$\bar{I}_{45} = I_0 + I_{90} - I_{135} \tag{3-41}$$

$$\bar{I}_{90} = -I_0 + I_{45} + I_{135} \tag{3-42}$$

$$\bar{I}_{135} = I_0 - I_{45} + I_{90} \tag{3-43}$$

Based on the above relationship we can estimate the given point information using the

neighborhood values of three different polarization directions.

（6）Inter-channel correlation

The Inter-channel correlation (ICC) method uses three steps of interpolation, noise reduction, and optimization to refine the demosaicking process. Only the interpolation part is discussed in this paper (Liu et al. 2020). The interpolation part draws on the CFA interpolation algorithm that utilizes inter-channel correlation, i.e., the luminance channel and the color channel are highly correlated. In the polarization image, the total intensity channel S_0 and the four polarization channels are highly correlated. Since there is a difference between the Bayer pattern of CFA and the mosaic pattern of polarization, the filter used in CFA needs to be adjusted before it can be used for polarization demosaicking. The adjusted convolution kernels are shown in Figure 3-5, where (a), (b) and (c) indicate the convolution kernels applied at 90°, 45° and 135° pixels when interpolating 0° channels respectively. The polarization angle channel where the points are located is marked in white and the weights are marked in black.

（a）   （b）   （c）
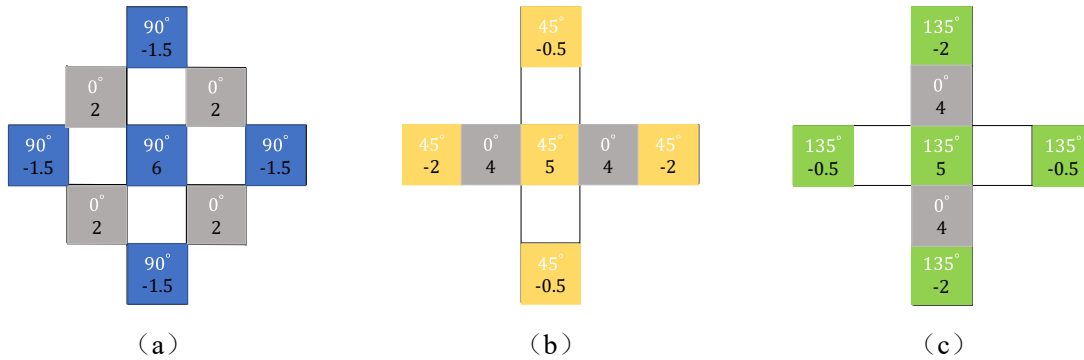
Fig. 3-5 Interpolated pixel weights

Demosaicking algorithm: Take the 0° channel recovery of Figure 3-5 as an example, each pixel at each position can be recovered by using its corresponding filter weighted sum to obtain the complete 0° channel recovery. The other three polarization channels are recovered in the same way.

3.2.2 Adaptive weighted interpolation

（1）Binary tree

The binary tree (BT) method uses weighted bilinear for interpolation, where the weights are determined by the inverse of the directional gradient (Miao et al. 2006; Mihoubi et al. 2018). This method was first applied in the multichannel demosaicking method, and its main idea is

to determine the channel interpolation order by the channel pixel distribution probabilities (starting from the one with the higher probability of interpolation and bifurcating downward in order, whose pixel distribution probabilities form a bifurcated tree diagram, and the frequency of occurrence of a channel sampling point can only be 2^n, so the method is called the bifurcated tree method) and the distribution pattern, and to transform its pixel distribution when interpolating a single channel. The interpolation process takes into account the inter-channel correlation, i.e., the weighted sum of the four neighborhood values, with the weighting factor determined by the 'edge value', and the larger the edge value, the smaller the weight, so as to ensure the maximum reproduction of the edges:
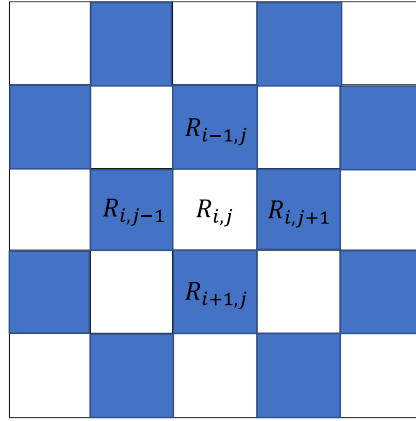
Fig. 3-6 Basic interpolation pattern

$$V: W_{m,n} = (1 + |R_{m+2,n} - R_{m,n}| + |R_{m-2,n} - R_{m,n}| + \\ 0.5|R_{m-1,n-1} - R_{m+1,n-1}| + 0.5|R_{m-1,n+1} - R_{m+1,n+1}|)^{-1} \qquad (3-44)$$

where $m \in \{i-1, \ i+1\}, n = j$.

$$H: W_{m,n} = (1 + |R_{m,n+2} - R_{m,n}| + |R_{m,n-2} - R_{m,n}| + \\ 0.5|R_{m+1,n-1} - R_{m+1,n+1}| + 0.5|R_{m-1,n-1} - R_{m-1,n+1}|)^{-1} \qquad (3-45)$$

where $m = i, n \in \{j-1, \ j+1\}$, R is the weighted average of the four neighboring points according to the above weights. However, there is no primary or secondary channel in a polarized image with only four channels, so the method is greatly simplified in terms of the interpolation sequence, first interpolating the pixel points located at the diagonal positions of known points, and each point is weighted and summed with the points in the four corners according to the diagonal gradient (analogous to the horizontal vertical gradient described above). Then the points at the remaining positions are estimated, and each point is weighted and summed by its four neighboring points according to the horizontal vertical gradient.

（2）Adaptive strategy

The Adaptive strategy (AS) proposes a locally adaptive 3×3 convolution kernel b, which is obtained by multiplying two convolution kernels correspondingly, one for the convolution kernel h_E determined from the Euclidean distance (the weight decreases with increasing distance) and one for the convolution kernel h_i determined from the intensity correlation under that neighborhood (the stronger the correlation the larger the coefficients), so as to ensures that the local interpolation does not cross the edges, thus preserving the high frequency components and reducing pseudo-edges as well as IFoV errors (Ratliff et al. 2011).

For each element surrounded by three different polarization pixels, h_E has three forms:

$$h_{E_v} = 0.5 \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{3-46}$$

$$h_{E_h} = 0.5 \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{3-47}$$

$$h_{E_d} = 0.25 \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \tag{3-48}$$

h_i is determined in this region in the following way:

$$h_i(k,l) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left\{-\frac{[I(m,n)-I(k,l)]^2}{2\sigma_i^2}\right\} \tag{3-49}$$

where σ_i is the standard deviation of the Gaussian penalty function, $k \in \{m-1, m, m+1\}$, $l \in \{n-1, n, n+1\}$. Then the expression of b is as follows:

$$b_v(m,n) = \frac{1}{a_v} \begin{bmatrix} 0 & h_{E_v}(1,2)h_i(1,2) & 0 \\ 0 & 0 & 0 \\ 0 & h_{E_v}(3,2)h_i(3,2) & 0 \end{bmatrix} \tag{3-50}$$

$$b_h(m,n) = \frac{1}{a_h} \begin{bmatrix} 0 & 0 & 0 \\ h_{E_h}(2,1)h_i(2,1) & 0 & h_{E_h}(2,3)h_i(2,3) \\ 0 & 0 & 0 \end{bmatrix} \tag{3-51}$$

$$b_d(m,n) = \frac{1}{a_d} \begin{bmatrix} h_{E_d}(1,1)h_i(1,1) & 0 & h_{E_d}(1,3)h_i(1,3) \\ 0 & 0 & 0 \\ h_{E_d}(3,1)h_i(3,1) & 0 & h_{E_d}(3,3)h_i(3,3) \end{bmatrix} \tag{3-52}$$

where

$$a_v = h_{E_v}(1,2)h_i(1,2) + h_{E_v}(3,2)h_i(3,2) \tag{3-53}$$

$$a_h = h_{E_h}(2,1)h_i(2,1) + h_{E_h}(2,3)h_i(2,3) \tag{3-54}$$

$$a_d = h_{E_d}(1,1)h_i(1,1) + h_{E_d}(1,3)h_i(1,3) + h_{E_d}(3,1)h_i(3,1) + h_{E_d}(3,3)h_i(3,3) \quad (3-55)$$

For each point of the original image, the four polarization information of the point can be recovered according to the above equation. Here we do not discuss the specific polarization angle channel, and take the recovered value of the channel with the relative interpolation point position in the diagonal D, vertical V, horizontal H, and center C (the point itself) as an example：

$$D(m,n) = b_d(1,1)I(m-1,n-1) + b_d(1,3)I(m-1,n+1) +$$
$$b_d(3,1)I(m+1,n-1) + b_d(3,3)I(m+1,n+1) \quad (3-56)$$

$$V(m,n) = b_v(1,2)I(m-1,n) + b_v(3,2)I(m+1,n) \quad (3-57)$$

$$H(m,n) = b_h(2,1)I(m,n-1) + b_h(2,3)I(m,n+1) \quad (3-58)$$

$$C(m,n) = I(m,n) \quad (3-59)$$

Due to the presence of

$$I_0 = I_{45} - I_{90} + I_{135} \quad (3-60)$$

$$I_{45} = I_0 + I_{90} - I_{135} \quad (3-61)$$

$$I_{90} = -I_0 + I_{45} + I_{135} \quad (3-62)$$

$$I_{135} = I_0 - I_{45} + I_{90} \quad (3-63)$$

Optimize the above D, V, H and C as

$$\widehat{D}(m,n) = H(m,n) + V(m,n) - C(m,n) \quad (3-64)$$

$$\widehat{V}(m,n) = C(m,n) + D(m,n) - H(m,n) \quad (3-65)$$

$$\widehat{H}(m,n) = C(m,n) + D(m,n) - V(m,n) \quad (3-66)$$

$$\widehat{C}(m,n) = C(m,n) \quad (3-67)$$

The four-channel recovery value of the point (m,n) can be obtained, and then the full plane can be recovered according to this method.

3.2.3 Edge detection interpolation

（1）Gradient-based

Gradient-based (GB) algorithm (Gao and Gruev 2013) uses sparse matrices to calculate the gradient of unknown points and determine the edges, which are subsequently interpolated along the edges using the bicubic method or bilinear method in smooth regions.

Definition of edge and calculation of gradient: a spatial gradient above a threshold is considered as an edge of the pixel point existing in some direction. Because in polarization imaging, edges indicate that two objects have different intensities, polarization angles or linear polarization, and these differences can be reflected in the four-channel image gradient. The algorithm defines the gradients in four directions (0°, 90°, 45° and 135°) for a pixel (i, j) using its 7×7 neighborhood and they can be computed in a sparse matrix. This convolution kernel size is experimentally derived to accurately detect edges and reduce the detected noise. If it is too small it will be sensitive to noise leading to detection of pseudo-edges, and too large it will smooth the real edges.

$$\begin{cases} D_{0°} = \sum_{i=2,4,6} \sum_{j=3,5,7} |I(i,j) - I(i,j-2)| \\ D_{45°} = \sum_{i=1,3,5} \sum_{j=3,5,7} |I(i,j) - I(i+2,j-2)| \\ D_{90°} = \sum_{i=3,5,7} \sum_{j=2,4,6} |I(i,j) - I(i-2,j)| \\ D_{135°} = \sum_{i=1,3,5} \sum_{j=1,3,5} |I(i,j) - I(i+2,j+2)| \end{cases} \qquad (3-68)$$

Interpolation process: For each sparse channel, first interpolate the points in the diagonal direction from the known points, and then interpolate the points in the other directions.

When interpolating the points in the diagonal direction position, $D_{45°}$ and $D_{135°}$ are obtained according to the method mentioned above

$$\frac{D_{45°}}{D_{135°}} \begin{cases} > T_{threshold}, & \text{Here there is an edge along } 135° \\ = T_{threshold}, & \text{Here there is an edge along } 45° \text{ or } 135° \\ < T_{threshold}, & \text{Here there is an edge along } 45° \end{cases} \qquad (3-69)$$

Bicubic interpolation is performed along the edges for cases with edges, otherwise bilinear interpolation is performed. The pixels at the remaining positions are interpolated on this basis, at which point $D_{0°}$ and $D_{90°}$ are calculated as described above.

$$\frac{D_{0°}}{D_{90°}} \begin{cases} > T_{threshold}, & \text{Here there is an edge along } 90° \\ = T_{threshold}, & \text{Here there is an edge along } 0° \text{ or } 90° \\ < T_{threshold}, & \text{Here there is an edge along } 0° \end{cases} \qquad (3-70)$$

Similarly, bicubic interpolation is performed along the edges for the case with edges, otherwise bilinear interpolation is performed.

Determination of $T_{threshold}$：Select different thresholds and calculate the RMSE of DoLP and AoLP after demosaicking, and select the threshold that minimizes the value, which was experimentally set to 60%.

（2）Novel smoothness based

This paper improves on the GB method, the GB algorithm is limited to diagonal, horizontal and vertical judgments for edges, but in fact edges can appear in any direction (Zhang et al. 2017). Calculating only four directional gradients is not accurate enough.

In this paper, instead of using gradient calculation to judge edges, we use the way of calculating smoothness. If the change of smoothness is greater than the threshold, bicubic interpolation is used, otherwise bilinear interpolation is used. First determine the threshold value, segment the original image according to 2×2 superpixels, and calculate the variance in each superpixel

$$J = \frac{1}{3}[(f(x,y) - E)^2 + (f(x+1,y) - E)^2 + \\ (f(x,y+1) - E)^2 + (f(x+1,y+1) - E)^2] \qquad (3-71)$$

$$\sigma = \sqrt{J} \qquad (3-72)$$

where E is the 4-element mean value in the superpixel.

The middle value σ_{median} of all superpixels σ is taken as the threshold value for judging the smoothness. If the superpixel $\sigma < \sigma_{median}$, the block is a smooth region and all three unknown quantities of the block are interpolated with bilinear interpolation, otherwise it is regarded as a non-smooth region and bicubic interpolation is used.

（3）Intensity correlation among polarization channels

The principle of intensity correlation among polarization channels (IC) to determine the edge based on inter-channel correlation is (Zhang et al., 2016), if a pixel point is a smooth point in a 3×3 area, all its polarization information can be obtained by averaging the polarization information of surrounding points. The operation is obtained:

$$\begin{pmatrix} \hat{I}_0 \\ \hat{I}_{45} \\ \hat{I}_{90} \\ \hat{I}_{135} \end{pmatrix} = \begin{bmatrix} 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & -1 \\ -1 & 1 & 0 & 1 \\ 1 & -1 & 1 & 0 \end{bmatrix} \begin{pmatrix} \bar{I}_0 \\ \bar{I}_{45} \\ \bar{I}_{90} \\ \bar{I}_{135} \end{pmatrix} \qquad (3-73)$$

where $\hat{I}_0$, $\hat{I}_{45}$, $\hat{I}_{90}$ and $\hat{I}_{135}$ are the estimated values and $\bar{I}_0$, $\bar{I}_{45}$, $\bar{I}_{90}$ and $\bar{I}_{135}$ are the mean

values obtained using the surrounding value estimates.

If the error between the polarization information of the pixel and the estimated value is large, the point is more likely to be an edge point, i.e., the more different from the polarization information of the neighboring points. This error can be expressed by the following equation to distinguish smooth regions, horizontal edges, vertical edges and diagonal edges.

$$\begin{cases} \Delta I^h(i,j) = |w_{H1}(i,j) \otimes I(i,j) - w_{H2}(i,j) \otimes I(i,j)| \\ \Delta I^v(i,j) = |w_{V1}(i,j) \otimes I(i,j) - w_{V2}(i,j) \otimes I(i,j)| \end{cases} \tag{3-74}$$

where $w_{V1} = w_{H1}{}^T, w_{V2} = w_{H2}{}^T$, $w_{H1} = \begin{bmatrix} -0.5 & 0.5 & -0.5 \\ 0.5 & -1 & 0.5 \\ 0 & 0.5 & 0 \end{bmatrix}, w_{H2} = \begin{bmatrix} 0 & 0.5 & 0 \\ 0.5 & -1 & 0.5 \\ -0.5 & 0.5 & -0.5 \end{bmatrix}$.

Calculate ΔI^h and ΔI^v for a pixel point. If the point is in a smooth area, they are equal, ΔI^h is 1 and ΔI^v is 0 for horizontal edges, ΔI^v is 1 and ΔI^h is 0 for vertical edges, and both are 0.5 for diagonal edges.

Interpolation process: For each sparse channel, the pixels located in the diagonal direction of the known point are interpolated first, and then the pixels located in the horizontal and vertical directions are interpolated. When interpolating diagonally located pixels, the gradient method is used to determine which kind of region the point is located in (smooth region, 45° edge or 135° edge) and the corresponding interpolation method is adopted. If $\Delta D_{45°} = \Delta D_{135°}$, the point is located in the smooth region and interpolated by taking the average of the four corners; if $\Delta D_{45°} > \Delta D_{135°}$, the point is located on the 135° edge and interpolated by bicubic spline interpolation along the 135° direction, otherwise the point is on the 45° edge and interpolated along that direction.

$$\begin{cases} \Delta D_{45°} = \sum_{k \in \{3,1,-1\}} |I(i+k,j-k) - I(i+k-2,j-k+2)| + A \\ \Delta D_{135°} = \sum_{k \in \{3,1,-1\}} |I(i+k,j+k) - I(i+k-2,j+k-2)| + B \end{cases} \tag{3-75}$$

where $A = |2I(i,j) - I(i-2,j+2) - I(i+2,j-2)|$, $B = |2I(i,j) - I(i-2,j-2) - I(i+2,j+2)|$.

Interpolation of pixels at horizontal and vertical positions: the error is calculated using the correlation method described above to determine whether the point is on a smooth, horizontal edge or vertical edge, calculated as follows:

$$h(i,j) = \sum_{k=-1}^{1} \sum_{t=-1}^{1} \Delta I^h(i+k,j+t) \tag{3-76}$$

$$v(i,j) = \sum_{k=-1}^{1} \sum_{t=-1}^{1} \Delta I^v(i+k, j+t) \qquad (3-77)$$

If $h(i,j) > v(i,j)$, the four adjacent values in the horizontal direction are interpolated in cubic spline; if $h(i,j) < v(i,j)$, the values are interpolated along the vertical direction; otherwise, the four corner values are averaged.

3.2.4 Difference and residual interpolation

（1）Residual interpolation

Residual interpolation (RI) first appeared in the CFA demosaicking algorithm (Kiku et al. 2013), but other literature (Mihoubi et al. 2018) has adapted it for application in polarization demosaicking.

For a certain sparse channel $\tilde{I}_k$, the 2018 literature uses the guided filtering method to generate the pre-estimated channel $\check{I}_k$ by up-sampling it in combination with the bootstrap image S_0. Then the residual channel is calculated, i.e., $\tilde{I}_k - \check{I}_k$ is calculated at the position of the original value of the k-channel to obtain the sparse matrix ΔI_k, and $\Delta \check{I}_k$ is obtained after bilinear interpolation of ΔI_k, then the final k-channel interpolation The result $\hat{I}_k = \check{I}_k + \Delta \check{I}_k$.

In the CFA algorithm, the guide image is the luminance channel, i.e., the green channel, while the S_0 channel derived from the original mosaic image is used as the guide image in the polarization algorithm since there is no main channel

$$S_0 = h * I_{raw} \qquad (3-78)$$

where

$$h = 0.5 \times \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Interpolation using residual channels is an evolution of difference interpolation between channels. Since the scenes obtained between different channels have similar details and edges, the differential sparse channel obtained by differencing can better erase the edges and get a smoother image that is easier to interpolate. The interpolation of the residual channel changes the two polarization channels to be differenced directly into one channel and another pre-estimated channel, so that the resulting residual channel will be smoother than the differential channel. The way to generate the pre-estimated channel is guided filter upsampling, and this method can preserve the scene edges well.

The guided filter upsampling process (He et al., 2012; Ahmed et al, 2017) is as follows:

the sparse matrix I_{intk} is upsampled using the guided image I_k. The guided filter is a linear model between the sparse channel and the output channel. The output value q_i is a linear transformation of the guided image in the neighborhood w_k centered at pixel k. The following is an example of a 0° channel:

$$q_{i0} = a_k I_i + b_k, \forall i \in w_k \qquad (3-79)$$

where a_k, b_k are the continuous linear coefficients under this window, which are found according to the principle of lowest cost function:

$$a_k = \frac{\frac{1}{w}\Sigma_{i\epsilon w_k} I_0 I_{int0} - \mu_k I_{int0k}}{\sigma_k^2 + \varepsilon} \qquad (3-80)$$

$$b_k = I_{int0k} - a_k \mu_k \qquad (3-81)$$

where μ_k and σ_k^2 are the mean and variance of the channel in window w_k respectively, w is the number of pixels, and I_{int0k} is the mean value of I_{int0}. Since each pixel can be computed with different coefficients under different windows, resulting in different mapping results, all possible values need to be averaged to obtain the final output at that point as follows:

$$\begin{cases} \overline{q_{i0}} = \frac{1}{w}\Sigma_{i\epsilon w_k}(a_k I_0 + b_k) \\ \overline{q_{i45}} = \frac{1}{w}\Sigma_{i\epsilon w_k}(a_k I_{45} + b_k) \\ \overline{q_{i90}} = \frac{1}{w}\Sigma_{i\epsilon w_k}(a_k I_{90} + b_k) \\ \overline{q_{i135}} = \frac{1}{w}\Sigma_{i\epsilon w_k}(a_k I_{135} + b_k) \end{cases} \qquad (3-82)$$

At this point, the process of upsampling the 0° sparse channel using the guided image S$_0$ is completed. This process can be implemented using the imguidedfilter function in MATLAB.

（2）Adaptive residual interpolation

Adaptive residual interpolation (ARI) improves the RI method by performing Laplace filtering on both the sparse channel and the bootstrap image followed by guided filtering upsampling, thus residual interpolation (Monno et al. 2015).

（3）Modified residual interpolation

Modified residual interpolation (MRI) is essentially the same as the RI method, except that instead of S_0, its guiding image corresponds to the full channel after bilinear interpolation for each sparse channel (Ahmed et al, 2017).

（4）Spectral difference

Spectral difference (SD) uses inter-channel correlation for differential channel interpolation to recover a blank point P in the k-channel as an example (a pixel on the p-channel in the original image), with the k-channel point contained within its superpixel set to K (Mihoubi et al. 2018; Brauers and Aach 2008).

Calculate $\Delta = \hat{I}_p - I_k$, where $\hat{I}_p$ is the value at point K after bilinear interpolation of the p-channel, and obtain $\hat{\Delta}$ after bilinear interpolation of Δ, then the k-channel interpolation at point P is $I_p - \hat{\Delta}_p$.

With this method, all the information at the missing positions can be recovered.

（5）Newton's polynomial and difference

In this paper, Newton's polynomial is used to interpolate in the difference channel (NPD) (Li et al. 2019). Since the error introduced by interpolation is usually more obvious in the high-frequency region, and the high-frequency parts such as edges and details are similar between different channels, this paper assumes that the channel difference can reduce the presence of high-frequency parts, and verifies the hypothesis by experiment, so as to improve the interpolation accuracy using Newton's interpolation in relatively smooth channels.

If there are $x_0, x_1, \dots, x_n$ discrete points in the closed interval $[a, b]$, the nth-order Newton polynomial and its error $R_n(x)$ are expressed as follows:

$$N_n(x) = f(x_0) + f[x_0, x_1](x - x_0) + \cdots +$$
$$f[x_0, x_1, \dots, x_n](x - x_0)(x - x_1) \dots (x - x_{n-1}) \qquad (3-83)$$

$$R_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(c) \prod_{i=0}^{n} (x - x_i) \qquad (3-84)$$

where $f[x_0, x_1, \dots, x_n]$ is the division difference.

In the interpolation method to solve the mosaic, take $n=1$, the two endpoints are x_0, $x_1 = x_0 + h$, and the middle point is $x = x_0 + 0.5h$; then $f(x)$ becomes

$$f\left(x_0 + \frac{h}{2}\right) \cong \frac{f(x_0 + h) + f(x_0)}{2} - \frac{f''(\xi)}{8} h^2 \qquad (3-85)$$

Taking the interpolated I_{90} channel as an example, let the centroid of a certain 7×7 neighborhood be $I_{90}(i, j)$, then $h=2$.

$$\tilde{I}_{90}(i, j-1) \cong \frac{I_{90}(i, j-2) + I_{90}(i, j)}{2} - \frac{f''(\xi)}{8} h^2 \qquad (3-86)$$

The second of these items is

$$\frac{f''^{(\xi)}}{8}h^2 \cong \frac{I_{90}(i,j+1) - 2I_{90}(i,j-1) + I_{90}(i,j-3)}{8} \qquad (3-87)$$

Since we do not know $I_{90}(i,j+1)$, $I_{90}(i,j-1)$, $I_{90}(i,j-3)$, we define the inter-channel difference Λ to represent the above equation

$$\Lambda = I_{135} - \hat{I}_{90} \qquad (3-88)$$

$$\frac{f''^{(\xi)}}{8}h^2 \cong \frac{I_{135}(i,j+1) - \Lambda(i,j+1)}{8} - \frac{I_{135}(i,j-1) - \Lambda(i,j-1)}{4} +$$
$$\frac{I_{135}(i,j-3) - \Lambda(i,j-3)}{8} \qquad (3-89)$$

Let the difference between channels agree in a small region: $\Lambda(i,j+1) \approx \Lambda(i,j-1) \approx \Lambda(i,j-3)$. Then the above equation can be abbreviated as

$$\frac{f''^{(\xi)}}{8}h^2 \cong \frac{I_{135}(i,j+1) - 2I_{135}(i,j-1) + I_{135}(i,j-3)}{8} \qquad (3-90)$$

Then $\tilde{I}_{90}(i,j-1)$ can be expressed in terms of known quantities as

$$\tilde{I}_{90}(i,j-1) \cong \frac{I_{90}(i,j-2) + I_{90}(i,j)}{2} -$$
$$\frac{I_{135}(i,j+1) - 2I_{135}(i,j-1) + I_{135}(i,j-3)}{8} \qquad (3-91)$$

$I_{90}(i,j+1)$ can be estimated in a similar way.

Since the difference channel still has a small amount of high-frequency parts, which are not completely smooth, an edge classifier is used to optimize the interpolation of the detail parts, and the following is an example of interpolating the diagonal direction elements such as (i,j) on 0° information. The horizontal and vertical direction is similar. In this 7×7 neighborhood, two I_0 estimates can be obtained from 45° and -45° diagonal interpolation, respectively, and the two estimates are weighted and summed.

$$\hat{I}_0^{45°}(i,j) = I_{90}(i,j) + \tilde{\Lambda}^{45°} \cong I_{90}(i,j) + \frac{\tilde{\Lambda}(i+1,j-1) + \tilde{\Lambda}(i-1,j+1)}{2} \qquad (3-92)$$

where $\tilde{\Lambda} = I_0 - I_{90}$. Using the polarization difference model it is possible to find

$$\hat{I}_0^{45°}(i,j) \cong I_{90}(i,j) + \frac{I_0(i+1,j-1) - \tilde{I}_{90}(i+1,j-1)}{2} +$$

$$\frac{I_0(i-1,j+1) - \tilde{I}_{90}(i-1,j+1)}{2}$$

$$= I_{90}(i,j) + \frac{I_0(i+1,j-1) + I_0(i-1,j+1)}{2} -$$

$$\frac{I_{90}(i+1,j-1) + \tilde{I}_{90}(i-1,j+1)}{2} \tag{3-93}$$

$$\hat{I}_0^{-45°}(i,j) \cong I_{90}(i,j) + \frac{I_0(i-1,j-1) - \tilde{I}_{90}(i-1,j-1)}{2} +$$

$$\frac{I_0(i+1,j+1) - \tilde{I}_{90}(i+1,j+1)}{2}$$

$$= I_{90}(i,j) + \frac{I_0(i-1,j-1) + I_0(i+1,j+1)}{2} -$$

$$\frac{I_{90}(i-1,j-1) + \tilde{I}_{90}(i+1,j+1)}{2} \tag{3-94}$$

where the third term can be derived by quoting the previously mentioned equation applied in the diagonal direction. Define the edge classifier according to the loss function ϕ

$$\varphi^{45°} = \sum_{m=\{-2,0,2\}} \sum_{n=\{-2,0,2\}} \left| \hat{I}_0^{45°}(i+m,j+n) - I_{90}(i+m,j+n) \right| \tag{3-95}$$

$$\varphi^{-45°} = \sum_{m=\{-2,0,2\}} \sum_{n=\{-2,0,2\}} \left| \hat{I}_0^{-45°}(i+m,j+n) - I_{90}(i+m,j+n) \right| \tag{3-96}$$

$$\phi = \max\left(\frac{\varphi^{45°}}{\varphi^{-45°}}, \frac{\varphi^{-45°}}{\varphi^{45°}} \right) \tag{3-97}$$

Then the estimated value of I_0 is

$$\hat{I}_0(i,j) = \begin{cases} \hat{I}_0^{45°}(i,j) & (\phi > \tau)(\phi = \dfrac{\varphi^{-45°}}{\varphi^{45°}}) \\ \hat{I}_0^{-45°}(i,j) & (\phi > \tau) \& (\phi = \dfrac{\varphi^{-45°}}{\varphi^{45°}}) \\ \dfrac{w^{45°}\hat{I}_0^{45°}(i,j) + w^{-45°}\hat{I}_0^{-45°}(i,j)}{w^{45°} + w^{-45°}} & (\phi \leq \tau) \end{cases} \tag{3-98}$$

where the weights w are

$$w^{45°} = \frac{1}{d^{45°} + \varepsilon} \tag{3-99}$$

$$w^{-45°} = \frac{1}{d^{-45°} + \varepsilon} \tag{3-100}$$

ε is a small quantity that ensures that the denominator is not zero, and to simplify the operation, d is the difference in the two directions:

$$
\begin{aligned}
d^{45°} = &|I_0(i+1,j-1) - I_0(i-1,j+1)| + \\
&|2I_{90}(i,j) - I_{90}(i-2,j+2) - I_{90}(i+2,j-2)|
\end{aligned} \qquad (3-101)
$$

$$
\begin{aligned}
d^{-45°} = &|I_0(i-1,j-1) - I_0(i+1,j+1)| + \\
&|2I_{90}(i,j) - I_{90}(i-2,j-2) - I_{90}(i+2,j+2)|
\end{aligned} \qquad (3-102)
$$

Next, consider applying the above classifier to interpolate the points at the horizontal and vertical position, such as to obtain the 45° and 135° polarization information at the position (i,j): when seeking the 45° polarization angle information, we can only interpolate along the vertical direction to obtain $\hat{I}_{45}^V(i,j)$, because there is no known 45° information in the horizontal direction, however, we can use the above method to interpolate the three values of this vacancy in turn to two diagonal. However, we can use the above-mentioned method to estimate these three vacant values in two diagonal directions, and then take the weighted average to obtain the trustable estimate, so that we can obtain the 45° polarization information $\hat{I}_{45}^H(i,j)$ obtained by interpolation in the horizontal direction, and then apply the above-mentioned edge classifier to the horizontal and vertical directions to obtain the final $\hat{I}_{45}(i,j)$ by weighting and summing the two. Similarly, we can obtain $\hat{I}_{135}(i,j)$.

After recovering all the points using the above method, the four polarization information of each point can be obtained, thus completing the mosaic solving process.

（6）Edge-aware residual

The Edge-aware residual interpolation (EAR) also uses residual interpolation after upsampling of the guided filter, except that unlike the guide images of the first two residual methods, this paper calculates a weighted S_0 as the guide image (Morimatsu et al. 2020).

The four-direction detection applied in this paper is inspired by horizontal and vertical edge detection (Zhang et al. 2016), and this paper proposes to estimate the intensity values of (i,j) by four directions: north, south, east and west, and determine whether there are horizontal, vertical or diagonal edges by the error between the actual value and the estimated value. Take the north direction n as an example, because S_0 has two expressions

$$
S_0 = I_0 + I_{90} = I_{45} + I_{135} \qquad (3-103)
$$

Therefore, two ways of estimating the intensity value S_0 at this point exist:

$$\widehat{S}_{0n(0,90)}(i,j) = I_0(i,j) + 0.5[I_{90}(i-1,j-1) + I_{90}(i+1,j-1)] \qquad (3-104)$$

$$\widehat{S}_{0n(45,135)}(i,j) = I_{45}(i,j-1) + 0.5[I_{135}(i-1,j) + I_{135}(i+1,j)] \qquad (3-105)$$

then there is

$$\widehat{S}_{0n}(i,j) = 0.5\left[\widehat{S}_{0n(0,90)}(i,j) + \widehat{S}_{0n(45,135)}(i,j)\right] \qquad (3-106)$$

$$\Delta\widehat{S}_{0n}(i,j) = \widehat{S}_{0n(0,90)}(i,j) - \widehat{S}_{0n(45,135)}(i,j) \qquad (3-107)$$

If the point is not at an edge, then $\Delta\widehat{S}_{0n}$=0. The estimate of $\widehat{S}_{0n}$ does not pass through the edge, so we evaluate the intensity difference between the four directions to determine the weight of the interpolation direction, and the larger the difference the smaller the weight, thus generating an edge-aware guided image.

Writing the above expression in convolutional form, S_0 under each direction should be

$$S_{0k} = F_k \otimes I_{PFA}, k = \{n, e, w, s\} \qquad (3-108)$$

$$F_n = \begin{bmatrix} 1/8 & 1/4 & 1/8 \\ 1/8 & 1/4 & 1/8 \\ 0 & 0 & 0 \end{bmatrix} \qquad (3-109)$$

$$F_e = \begin{bmatrix} 0 & 1/8 & 1/8 \\ 0 & 1/4 & 1/4 \\ 0 & 1/8 & 1/8 \end{bmatrix} \qquad (3-110)$$

$$F_w = \begin{bmatrix} 1/8 & 1/8 & 0 \\ 1/4 & 1/4 & 0 \\ 1/8 & 1/8 & 0 \end{bmatrix} \qquad (3-111)$$

$$F_s = \begin{bmatrix} 0 & 0 & 0 \\ 1/8 & 1/4 & 1/8 \\ 1/8 & 1/4 & 1/8 \end{bmatrix} \qquad (3-112)$$

$\Delta\widehat{S}_{0k}$ should be:

$$\Delta\widehat{S}_{0k} = H_k \otimes I_{PFA}, k = \{n, e, w, s\} \qquad (3-113)$$

$$H_n = \begin{bmatrix} -1/2 & 1 & -1/2 \\ 1/2 & -1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix} \qquad (3-114)$$

$$H_e = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 0 & -1 & 1 \\ 0 & 1/2 & -1/2 \end{bmatrix} \qquad (3-115)$$

$$H_w = \begin{bmatrix} -1/2 & 1/2 & 0 \\ 1 & -1 & 0 \\ -1/2 & 1/2 & 0 \end{bmatrix} \qquad (3-116)$$

$$H_s = \begin{bmatrix} 0 & 0 & 0 \\ 1/2 & -1 & 1/2 \\ -1/2 & 1 & -1/2 \end{bmatrix} \qquad (3-117)$$

Then the weights are

$$W_k(i,j) = \frac{1}{\Delta \widehat{S_0}'_k + \varepsilon} \qquad (3-118)$$

$$\Delta \widehat{S_0}'_k = M_k \otimes \left| \Delta \widehat{S_0}_k \right| \qquad (3-119)$$

where ε is a small amount (set to 10^{-32}) that guarantees that the denominator is not zero and M_k is a 5×5 smoothing filter. Then the weighted guide image G is expressed as:

$$G(i,j) = \sum_{k=n,e,w,s} W_k(i,j)X_k(i,j) \Big/ \sum_{k=n,e,w,s} W_k(i,j) \qquad (3-120)$$

Each sparse matrix was then upsampled and residuals interpolated using guided filter in the same process as the residuals interpolation method described above (Kiku et al. 2016).

（7）Difference prior

The difference prior (DP) method is distinct from the aforementioned residual-based methods. Instead of using residuals, DP directly completes interpolation based on the differences between channels. For a particular channel, the differences between it and the other three channels are calculated. After interpolation, these differences are added back to obtain three estimated matrices for that channel. Leveraging inter-channel correlations, these three results are then weighted and summed to derive the final estimated result for that channel (Wu et al. 2021).

Before taking the differences, bilinear interpolation is applied to the other three sparse matrices to prepare for the differences at corresponding points to the sparse channel. The difference channel is subjected to bilinear interpolation once again to obtain the complete difference channel. The subtracted channel is then added back to yield three estimated matrices for that channel. Based on research into inter-channel correlations, coefficients of correlation are determined for different channels. These coefficients are used for weighted summation according to the calculated correlations to derive the final estimated result for that channel. The following section focuses on how to derive these coefficients through the study of inter-channel correlations.

Differences between polarized channels create the difference channel. When the

polarization angles of two channels are orthogonal, the differences are more significant, making the edges of the difference channel more pronounced compared to non-orthogonal configurations. Thus, non-orthogonal channel correlations are stronger than orthogonal ones. This phenomenon is experimentally and theoretically proven, with quantitative calculations provided in the literature. Using a colored polarized image database (Qiu et al. 2019), images from the green channel are taken, and the differences between channel I_0 and the other three channels are studied. Distinct edge characteristics in different edge regions of various difference channels under different illumination sources are observed, with the most prominent edges found in the $I_0 - I_{90}$ group.

Theoretical derivations yield correlation coefficients between different channels. If the intensity of incident light I is regarded as a combination of non-polarized and polarized components, and considering Malus's law, it is concluded that

$$I_0 = I_{linear}\cos^2\varphi + \frac{I_{unpolar}}{2} \qquad (3-121)$$

$$I_{45} = I_{linear}\cos^2\left(\frac{\pi}{4} - \varphi\right) + \frac{I_{unpolar}}{2} \qquad (3-122)$$

$$I_{90} = I_{linear}\cos^2\left(\frac{\pi}{2} - \varphi\right) + \frac{I_{unpolar}}{2} \qquad (3-123)$$

$$I_{135} = I_{linear}\cos^2\left(\frac{3\pi}{4} - \varphi\right) + \frac{I_{unpolar}}{2} \qquad (3-124)$$

Where φ is the linear polarization angle. The difference between the I_0 channel and the other channels is

$$\Delta^{0,45} = I_0 - I_{45} = \frac{I_{linear}(\cos(2\varphi) - \sin(2\varphi))}{2} \qquad (3-125)$$

$$\Delta^{0,90} = I_0 - I_{90} = I_{linear}\cos(2\varphi) \qquad (3-126)$$

$$\Delta^{0,135} = I_0 - I_{135} = \frac{I_{linear}(\cos(2\varphi) + \sin(2\varphi))}{2} \qquad (3-127)$$

The magnitude of this difference depends on the intensity of incident linearly polarized light and its polarization angle. Considering two adjacent pixels, p and q, the difference between the difference channels of these two pixels can be obtained by subtracting:

$$\sigma^{0,45} = \left|\Delta^{0,45}{}_p - \Delta^{0,45}{}_q\right|$$
$$= \frac{\left|I_p(\cos(2\varphi_p) - \sin(2\varphi_p)) - I_q(\cos(2\varphi_q) - \sin(2\varphi_q))\right|}{2}$$

$$(3-128)$$

$$\sigma^{0,90} = \left| \Delta^{0,90}{}_p - \Delta^{0,90}{}_q \right| = \left| I_p \cos(2\varphi_p) - I_q \cos(2\varphi_q) \right| \qquad (3-129)$$

$$\sigma^{0,135} = \left| \Delta^{0,135}{}_p - \Delta^{0,135}{}_q \right|$$

$$= \frac{\left| I_p \big(\cos(2\varphi_p) + \sin(2\varphi_p) \big) - I_q \big(\cos(2\varphi_q) + \sin(2\varphi_q) \big) \right|}{2}$$

$$(3-130)$$

Assuming that linearly polarized light is uniformly distributed throughout the entire image, the global difference can be calculated using the integration of σ, providing a quantitative description of the entire difference channel:

$$\Sigma^{0,45} = \int_0^1 \int_0^1 \int_0^\pi \int_0^\pi \sigma^{0,45} dI_p dI_q d\varphi_p d\varphi_q \qquad (3-131)$$

$$\Sigma^{0,90} = \int_0^1 \int_0^1 \int_0^\pi \int_0^\pi \sigma^{0,90} dI_p dI_q d\varphi_p d\varphi_q \qquad (3-132)$$

$$\Sigma^{0,135} = \int_0^1 \int_0^1 \int_0^\pi \int_0^\pi \sigma^{0,135} dI_p dI_q d\varphi_p d\varphi_q \qquad (3-133)$$

The analytical solution of the above three equations is too complicated, but its ratio can be found:

$$\Sigma^{0,45} : \Sigma^{0,90} : \Sigma^{0,135} = 1 : \sqrt{2} : 1 \qquad (3-134)$$

The same is true for the ratio of the difference between the other channels.

Using a database of images to validate the aforementioned ratio, which includes 20 scenes with non-polarized illumination and 10 scenes with polarized illumination. The average distance between adjacent points, denoted as d_{ave}, is employed to measure the magnitude of inter-channel differences.

$$d_{ave} = \frac{1}{8(M-2)(N-2)S} \sum_{scenes=1}^{S} \sum_{i=2}^{M-1} \sum_{j=2}^{N-1} \sum_{m,n=(-1,0,1)} \left| P_{center}^{(i,j)} - P_{neighbor}^{(i,j)} \right| \qquad (3-135)$$

Here, M and N represent the width and length of the image, respectively. For polarized illumination scenes, $S = 10$, and for non-polarized illumination scenes, $S = 20$. The data indicates that in non-polarized illumination scenes, the d_{ave} ratio is approximately 1:1.2:1, while in scenes with polarized illumination, the off-diagonal difference channels differ slightly from the theoretical ratio.

This inconsistency may be related to the background in the images. The previous ratio was derived under the assumption of uniform distribution of linearly polarized light and AoLP in

the neighborhood. However, in real-world scenes, the background often contains large regions of uniform AoLP and polarization intensity, which disrupt this assumption and lead to a ratio closer to 1:1:1. Although this ratio differs from the theoretical one, it does not affect the interpolation method. The background region only contains low-frequency information and does not impact the weighted average calculation dominated by high-frequency information. Therefore, the final result I_i of a certain channel i should be weighted according to $1:\sqrt{2}:1$ relative to the difference channel j.

$$I_i = \sum w_{i,j}(\tilde{I}_j + \hat{\Delta}^{i,j}) \tag{3-136}$$

$$w_{i,j} = \begin{cases} \dfrac{\sqrt{2}}{1+\sqrt{2}+1} & , & if\,|i-j| \neq 90° \\ \dfrac{1}{1+\sqrt{2}+1} & , & if\,|i-j| = 90° \end{cases} \tag{3-137}$$

where $\tilde{I}_j$ represents the result of original data bilinear interpolation for channel j, and $\hat{\Delta}^{i,j}$ denotes the result of interpolation for the difference channel.

（8）Pseudo-panchromatic image difference

The psuedo-panchromatic image difference (PPID) method defines a psuedo-panchromatic image (PPI) and makes it interpolate the residuals with each of the four sparse channels, which are interpolated using a binary tree method (Mihoubi et al. 2015, 2016).

PPI is defined at each pixel as the mean value across all channels, and it has been demonstrated that PPI is highly correlated with all channels. Therefore, in the residual method, PPI can serve as an effective control channel for differencing. Given the strong spatial correlation between neighboring pixels (spectral correlation coefficient ≥ 0.9), the estimated PPI can be obtained from the original mosaic image using a convolution.

$$M = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \tag{3-138}$$

Hence, $PPI = I'_{raw} * M$, where I'_{raw} represents the image after luminance normalization of the original raw image I_{raw}:

$$I'_{raw} = I_{raw} .* \frac{\max(I_{raw})}{\max(I^k_{raw})} \tag{3-139}$$

This method is a simplification of the SFA demosaicking algorithm. In the original approach, further optimization was conducted by integrating directional gradients to enhance

the PPI edges. However, this step is not applicable to polarized images with only four channels, thus this step is omitted (Mihoubi et al. 2018).

（9）Edge compensation on correlation

The Edge Compensation on Correlation (ECC) method based on DoFP polarized image inter-channel correlation (Xin et al., 2023) was originally devised for CFA images. It utilized second-order differentiation for edge compensation, resulting in favorable demosaicking effects. Due to the limited edge detection involved, it exhibited higher computational efficiency. A typical algorithm used in the context was the Hamilton-Adams approach (Adams and Hamilton, 2018), which employed first-order and second-order derivatives to determine edge directions and guide interpolation. The ECC method adapted the edge compensation principle to DoFP images while simplifying the process of the Hamilton-Adams algorithm. The specific operations are outlined as follows:

[1] Begin by performing bilinear interpolation on each channel. Let the sparse matrix for channel k before interpolation be denoted as i_k , and after interpolation as i'_k;

[2] Calculate the second-order derivatives in both the horizontal and vertical directions for each channel. For example, the expression for channel k is as follows:

$$\partial_h^2 i'_k = \frac{i'_k(i, j + 2) + i'_k(i, j - 2) - 2i'_k(i, j)}{4} \qquad (3-140)$$

$$\partial_v^2 i'_k = \frac{i'_k(i + 2, j) + i'_k(i - 2, j) - 2i'_k(i, j)}{4} \qquad (3-141)$$

[3] Calculate the average of horizontally and vertically adjacent pixels:

$$\overline{i'_k}h = 0.5\left(i'_k(i, j + 1) + i'_k(i, j - 1)\right) \qquad (3-142)$$

$$\overline{i'_k}v = 0.5\left(i'_k(i + 1, j) + i'_k(i - 1, j)\right) \qquad (3-143)$$

[4] The preliminary estimation matrix for the 0-degree channel is given by:

$$\widehat{i_k} = 0.5\left[\left(\overline{i'_k}h - \partial_h^2 i'_k\right) + \left(\overline{i'_k}v - \partial_v^2 i'_k\right)\right] \qquad (3-144)$$

Channel Inter-correlation: In DoFP images, interpolation in difference image can effectively reduce interpolation errors caused by high-frequency energy (Zhao et al., 2019). By analyzing the correlation between the four polarized channels, this study determined a model for determining the difference image. Through an analysis of near-infrared polarized images acquired under different scenes and materials (Popescu et al., 2018), it was found that linear

relationships between channels did not directly exist. However, employing the Spearman rank-order correlation coefficient (SRC) revealed significant correlations between the polarized image channels regardless of the degree of polarization in different image regions. Notably, the correlations between diagonally opposite channels were weaker than those between non-diagonal channels. The Spearman rank-order correlation coefficient is defined as follows:

$$SRC(C_i, C_j) = \frac{\sum\left((I_i^p - \mu_i)(I_j^p - \mu_j)\right)}{\sqrt{\sum(I_i^p - \mu_i)^2}\sqrt{\sum\left(I_j^p - \mu_j\right)^2}} \qquad (3-145)$$

where C_i and C_j represent different polarization channels，I_i^p, I_j^p are the pixel values of channel i, j respectively，μ_i, μ_j are the means of the two channels.

Difference Model: Exploiting the strong inter-channel correlations, interpolation can be conducted with mutual reference between the four channels. Based on the aforementioned correlation study, it can be inferred that in high-frequency regions, the difference in diagonal channel interpolation is larger than that between non-diagonal channels (Zhao et al., 2019). Using diagonal differences for interpolation could lead to larger errors. Therefore, interpolation is only performed between non-diagonal channels.

Interpolation of inter-channel differences based on edge compensation：

[1] The result $\widehat{\imath_k}$ is obtained after the preliminary interpolation of channel k;

[2] The difference between its two non-diagonal channels m and n, and the sampled point of channel k is given by:

$$\Delta_{m,\ k} = \hat{\imath}_m \odot mask_k - i_k \qquad (3-146)$$

$$\Delta_{n,\ k} = \hat{\imath}_n \odot mask_k - i_k \qquad (3-147)$$

[3] The bilinear interpolation of the difference channels yields $\hat{\Delta}_{m,k}$，$\hat{\Delta}_{n,k}$;

[4] Using the two difference channels, combined with the subtracted channels m and n, obtain the two subtracted channels for k:

$$\tilde{\imath}_k^m = \hat{\imath}_m - \hat{\Delta}_{m,k} \qquad (3-148)$$

$$\tilde{\imath}_k^n = \hat{\imath}_n - \hat{\Delta}_{n,k} \qquad (3-149)$$

[5] These two channels represent the estimated values for channel k. However, as they are derived from different channels that are strongly correlated, the final demosaicked image for channel k is determined by taking their average (although the weights between the two

channels should ideally be calculated based on the specific conditions of the image, for simplification purposes, both weights are set to 0.5).

$$\tilde{\imath}_k = 0.5(\tilde{\imath}_k^m + \tilde{\imath}_k^n) \qquad\qquad (3-150)$$

3.2.5 Frequency domain replacement

（1）Discrete Wavelet Transform

The image is transformed into the frequency domain using discrete wavelet transform (DWT), and each channel is considered in terms of its low-frequency and high-frequency components. The demosaicking process exploits the strong correlation in the high-frequency components (Wang et al., 2015; Mihoubi et al., 2018). Initially, each sparse matrix is preliminarily interpolated using bilinear interpolation. In the frequency domain, the high-frequency components of all channels are replaced, while the low-frequency components are assumed to be adequately handled and not further optimized. The result is then transformed back into the spatial domain to obtain the final demosaicked result.

The specific replacement process is as follows: The reconstructed channel $\hat{I}_k$ can be divided into four down-sampled images (DSI), each consisting of pixels from the same polarized angle sampling point in the original image. Among them, one DSI (with polarization angle k) is entirely composed of actual sampled values and is referred to as DSI_k.

Perform a Haar wavelet transform on each DSI, and replace all the high-frequency bands of each channel's DSI_i except DSI_k with the high-frequency bands of the DSI_i corresponding to the 90-degree channel.

This algorithm was initially applied in the context of the SFA demosaicking algorithm, assuming that the mid-band channel's high-frequency part was the clearest and therefore suitable for replacement. However, in the case of four-channel polarized demosaicking, this condition does not exist. Therefore, any channel, such as I_{90}, can be arbitrarily chosen for replacement.

（2）PPI based DWT

The fundamental idea of the PPI-based Discrete Wavelet Transform (PPIDWT) method for pan-sharpening is to leverage the strong correlation that PPI exhibits with all channels. In this approach, the replacement criterion in the DWT method, which was originally I_{90}, is modified to use PPI instead. The remaining steps of the process remain the same as in the

standard DWT method (Mihoubi et al. 2016, 2018).

3.3 Summary of this chapter

This chapter begins by presenting conclusions drawn from statistical analysis of real image data, highlighting the strong inter-channel and intra-channel correlations. These conclusions serve as innovative insights for subsequent interpolation methods. The chapter then proceeds to provide a comprehensive overview of 22 typical polarized image demosaicking algorithms. For each algorithm, details are given including its characteristics, relationships with other algorithms, principles, and mathematical derivations. These algorithms are categorized into five main classes based on their core approaches: Low-pass filter interpolation, Adaptive weighted interpolation, Edge detection methods, Residual methods, and Frequency domain replacement methods. This categorization aids in understanding their main features, facilitating subsequent comprehensive evaluations for this category of algorithms.

Chapter 4 Computer Simulation Experiments and Algorithm Evaluation

After presenting the theoretical feasibility of each algorithm, this chapter implements the mathematical processes of the aforementioned algorithms using MATLAB. The algorithms are applied to publicly available image databases through network resources. Finally, the accuracy of algorithmic demosaicking is qualitatively and quantitatively evaluated using relevant metrics. The following section outlines the key aspects of the database used, highlights the essential steps of the computer simulation demosaicking process, and discusses the obtained data and image results from the experiments, followed by a comparative analysis.

4.1 Experiment

To compare the effectiveness and advantages and disadvantages of the above 22 interpolation demosaicking algorithms, we selected images from a publicly available database for testing, and each algorithm recovered the main polarization information and compared its PSNR.

The database used in this experiment consists of 12-channel color polarized images captured by Miki et al. It comprises 40 scenes, each containing color polarized images at 0°, 45°, 90°, and 135° polarization angles. These images have a bit depth of 10 bits and were captured under the same conditions by averaging 1000 shots to reduce random noise, effectively approximating noise-free real images (Morimatsu et al. 2020).

For this experiment, the first 30 scenes from this database were selected. Only the green channel of each RGB image was used as the ground truth for polarized information. After applying a mosaic pattern following the polarized image pattern in Figure 1-1 (b), the 22 demosaicking algorithms were applied to each image. The final outcome is the average PSNR between the demosaicked polarized information and the ground truth. The polarized information considered in this study includes I_0, I_{45}, I_{90}, I_{135}, S_0, S_1, S_2, DoLP and AoLP.

4.2 Analysis of results

4.2.1 For the analysis of PSNR results in this paper

Among the aforementioned 22 algorithms, some were accompanied by codes provided by the original authors. We made slight adjustments to these codes for execution. For the references where the original code was not provided, we created MATLAB programs based on their

descriptions. As a result, we computed the average PSNR values for each algorithm across 30 scenes and various polarized information types. The results are reported with four significant digits, as presented in Table 4-1：

Table 4-1 PSNR among different interpolation methods for different polarization information

method	I_0	I_{45}	I_{90}	I_{135}	S_0	S_1	S_2	DoLP	AoLP
Nearest	31.62	31.25	31.43	31.41	34.59	34.75	34.55	15.58	16.24
Bilinear	36.46	36.04	36.37	36.09	38.84	40.29	39.78	19.78	18.24
Bicubic	33.55	33.10	33.50	33.13	36.98	36.23	36.06	15.94	16.94
BS	29.51	30.92	36.95	31.73	33.99	35.72	35.31	15.38	16.65
B0	27.52	29.43	36.12	30.28	34.59	32.06	32.93	15.58	10.68
B1	35.82	32.86	39.20	31.83	37.02	32.87	34.63	18.02	10.53
B2	35.82	32.86	31.32	31.83	38.84	34.59	34.63	19.78	10.31
B3	30.06	30.71	31.18	30.84	33.68	36.14	35.49	21.87	10.29
B4	30.92	30.46	30.35	30.33	33.34	35.88	34.97	20.91	10.26
ICC	40.22	39.43	40.39	39.25	42.90	44.88	43.05	20.79	19.65
BT	36.22	35.82	36.14	35.87	37.85	41.71	41.01	21.18	18.74
AS	30.11	30.49	30.41	29.45	38.84	33.15	32.49	19.78	7.31
GB	35.66	35.30	35.60	35.42	38.09	39.68	39.33	18.72	17.96
NSB	29.98	31.54	36.86	32.35	35.90	35.98	35.82	17.71	17.25
IC	35.91	35.59	35.85	35.64	38.64	39.62	39.24	18.91	17.93
RI	33.70	33.38	33.61	33.41	35.35	38.50	38.44	18.12	15.89
ARI	38.21	37.71	38.20	37.86	40.88	41.84	41.30	20.56	18.76
MRI	33.42	33.08	33.37	33.11	35.14	38.10	38.01	17.82	15.84
SD	40.50	39.54	40.73	39.38	43.38	44.09	42.49	19.68	19.29
NPD	34.07	41.29	34.12	41.18	45.74	34.69	44.15	22.89	11.30
EAR	41.78	40.67	41.88	40.48	44.19	45.91	44.05	23.26	20.26
DP	40.94	39.88	40.95	39.83	43.27	45.09	43.52	21.93	19.96
PPID	39.43	38.13	38.07	38.34	40.31	43.10	42.49	21.52	18.76
ECC	39.84	39.23	39.82	39.17	42.39	43.35	42.47	21.48	19.35
DWT	36.46	36.80	32.72	36.47	38.73	38.10	41.49	18.24	14.61
PPIDWT	36.46	37.96	32.86	37.93	38.38	39.69	42.20	18.85	14.84

After comparison, it can be observed that both the Newton polynomial interpolation and the edge aware residual interpolation method exhibit the highest PSNR values. Based on this metric, these two algorithms are considered to provide more reliable and accurate results for polarized mosaic restoration. Specifically, the Newton polynomial interpolation achieves the highest signal-to-noise ratio in the I_{45}, I_{135}, S_0 and S_2 polarized image restorations, while the edge aware residual interpolation achieves the highest PSNR values in the I_0, I_{90}, S_1, AoLP and DoLP polarized image restorations.

Among the five major categories of interpolation methods, the residual-based methods generally achieve relatively higher PSNR values. Specifically, the adaptive residual interpolation, edge compensation on correlation method, spectral difference method, difference

prior method, and the pseudo-panchromatic image difference method all demonstrate PSNR values that are just slightly below those of the Newton polynomial interpolation and the edge-aware residual interpolation. This is attributed to the fact that most residual-based methods not only exploit the concept of channel-wise differencing but also incorporate edge detection, enhancing the accuracy of interpolation compared to direct edge detection during single-channel interpolation. The adaptive interpolation follows a similar trend.

Among the various spatial-invariant linear low-pass filtering methods, the channel-wise convolution with designed kernel method exhibits the best performance, ranking just below the Newton polynomial interpolation and the edge aware residual interpolation. This indicates that the analogy between CFA channel relationships and PFA methods is appropriate. Due to its incorporation of multi-channel relationships, this method outperforms both the other single-channel interpolation filtering methods and the bilinear with 5 kernels interpolation method based on point distance. Aside from the channel-wise convolution with designed kernel method, the classic bilinear interpolation achieves the best PSNR in almost all polarized information types (except for DoLP, where the B3 convolution kernel performs the best). Therefore, the bilinear interpolation is favored as a preprocessing step in many algorithms, especially when speed and memory efficiency are prioritized over high precision.

The performance of frequency domain replacement method is nearly identical to that of the bilinear interpolation, and it does not demonstrate notably exceptional results. Moreover, its computational process is more complex due to involving frequency domain transformations. The lack of significantly improved performance could be attributed to its higher demands on the distribution of polarized information, image smoothness, and noise level in the database scenes. This result may also be influenced by an insufficient number of samples or inappropriate scenarios.

Furthermore, research (LeMaster and Hirakawa 2014) has indicated that different mosaic polarization filter patterns can also impact the effectiveness of mosaic removal methods. However, this aspect is not extensively discussed in this paper. The study also refrains from analyzing the computational complexity or runtime of different algorithms, as MATLAB code execution time may not accurately represent algorithm complexity and cannot reflect their performance on specific polarized imaging hardware devices.

4.2.2 Analysis of results compared to other literature

Over the years, various algorithms have been proposed for the removal of mosaic artifacts in polarized images, but there are not many comprehensive articles summarizing such algorithms. In 2018, a review published in *Sensors* (Mihoubi et al. 2018) provided a thorough compilation and organization of these algorithms, hereafter referred to as the 'Survey'. The Survey summarized and explained a total of 19 different algorithms, including those adjusted from CFA and SFA, as well as algorithms directly applied to PFA, and compared them using the PSNR as the evaluation metric. Among the 22 algorithms studied in this article, 16 were mentioned in the Survey, and several typical mosaic artifact removal algorithms for polarized images that were not covered in the pre-2018 Survey and emerged after 2018 were also included. The following sections present a comparative analysis between the findings of this article and the Survey.

In the Survey's result analysis, the residual interpolation exhibited the highest PSNR values in S_2 and DOLP, while the pseudo-panchromatic image difference method achieved the highest PSNR values in other metrics. Although the results of this article differ slightly, the following analysis addresses both the differences and similarities.

This article asserts that the adaptive residual interpolation method performs significantly better across all polarization metrics compared to both the residual interpolation method and the improved residual interpolation. In contrast, the Survey suggests that the residual interpolation and the adaptive residual interpolation yield similar effects (with PSNR differences fluctuating between 0.09 and 1.39 across all metrics), while the residual interpolation slightly outperforms in S_2 and DoLP. This disparity can be attributed to differences in database selection, variations in parameters during the adjustment of CFA algorithms, and discrepancies in code processing details. Consequently, before 2018, this article identifies the adaptive residual interpolation, pseudo-panchromatic image difference, and spectral difference method as the preferable algorithms, deviating somewhat from the Survey's conclusions. Nevertheless, both sources concur that utilizing residual-based interpolation along with optimizations, such as introducing panchromatic imagery, leads to favorable interpolation outcomes.

Post-2018, two algorithms, the Newton polynomial method and the edge aware residual interpolation method, have emerged, both applying inter-channel interpolation and achieving

the most optimal results to date. Additionally, the difference prior method and the Edge compensation method based on inter-channel correlations also perform remarkably well, trailing slightly behind the Newton polynomial interpolation and the edge aware residual interpolation. Thus, it is evident that employing inter-channel differencing (or residual-based) interpolation and leveraging inter-channel correlations represent a promising approach for obtaining optimal solutions within classical interpolation methods. This insight can guide future innovations and enhancements in polarized mosaic artifact removal algorithms.
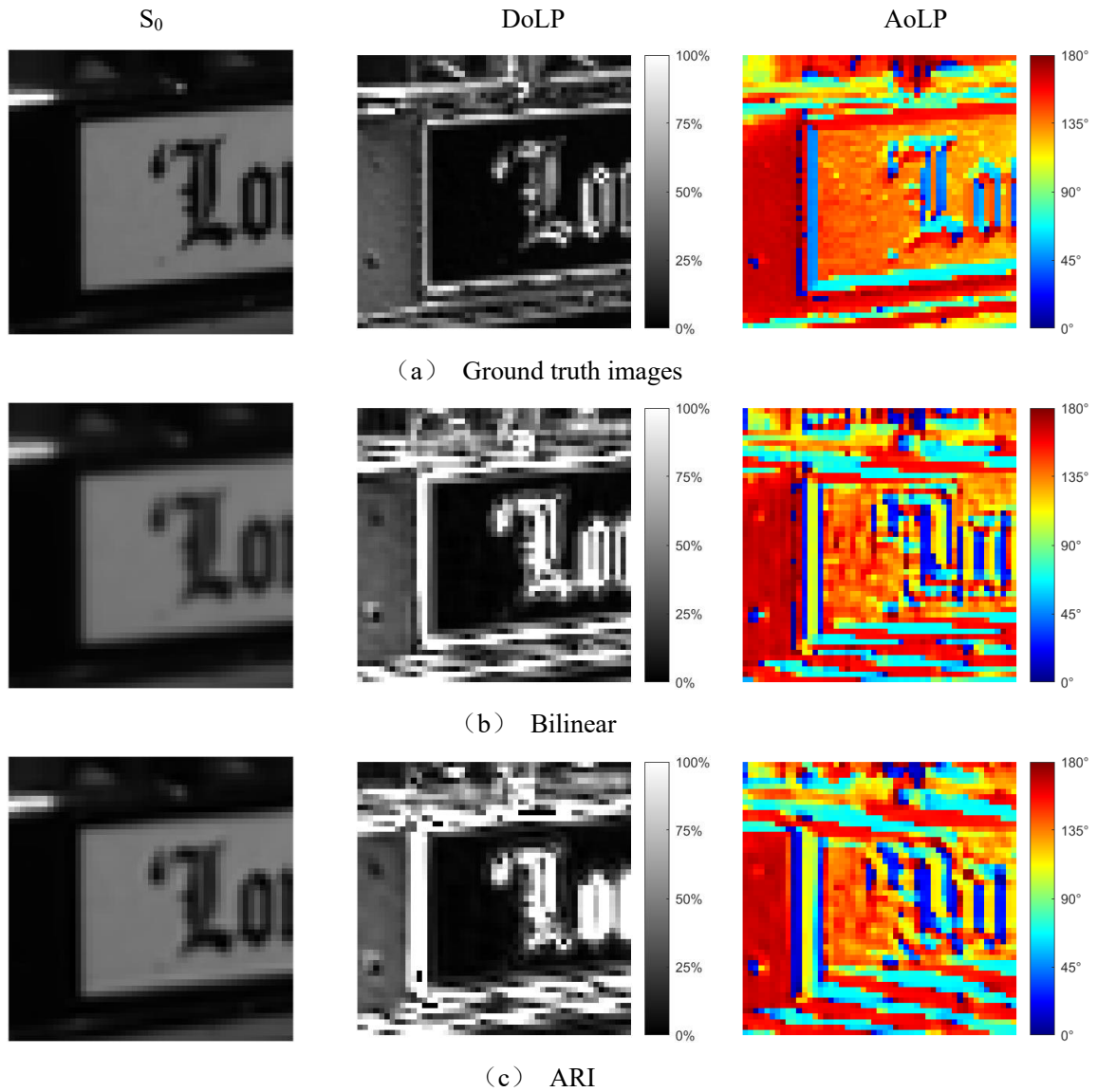
4.2.3 Subjective evaluation

The following is a subjective evaluation of several above-mentioned demosaicking algorithms with high PSNR index, including the classical bilinear method, adaptive residual interpolation, psuedo-panchromatic image interpolation, Newton polynomial interpolation, edge-aware residual interpolation, and comparison of the differences between their S_0, DoLP, AoLP and real images. Usually, since AoLP and DoLP are not obtained by direct demosaicking, but by Stokes vector calculation, they will be introduced with large errors, such as false edge, zipper noise, etc. Therefore, the examination of these two recovered images can often reflect the quality of the algorithms more directly. The examination scene is "bus", which shows the difference between the recovered image and the true value for a particular detail in the scene by different algorithms. As shown in Figure 4-1, each row is the recovered image of one algorithm, each column is the type of polarization information, and the first row is the actual value image.

For the S_0 image, all five methods recovered better, and both the bilinear and psuedo-panchromatic image difference methods had some smoothing effect at the edges, compared to the Newton polynomial interpolation method and the residual interpolation with edge knowledge method which were relatively clear at the edges.

For DoLP, the bilinear interpolation method obviously produces false edges at letters and borders, as do the adaptive residual interpolation method, the pseudo-panchromatic image difference method, and the residual interpolation method with edge knowledge, while the Newton interpolation method has less edge error in comparison, and the results obtained by the Newton interpolation method are the closest to the original image in terms of polarization values (the image details are close in gray scale).

For the case of AoLP (angle of linear polarization) images, several algorithms exhibit

significant degradation in performance. This is attributed to the fact that AoLP is calculated using the Stokes matrix combined with inverse trigonometric functions. Considering error propagation functions, its noise and errors are often substantial. Among these algorithms, bilinear interpolation and the adaptive residual interpolation method show prominent pseudo-edges. The Newton interpolation method yields less accurate AoLP values in this context. The most effective methods appear to be the edge knowledge residual interpolation method and the panchromatic image difference method, with the latter exhibiting evident zipper noise at edges. Consequently, the edge aware residual interpolation method seems to yield the best AoLP restoration results.

（a） Ground truth images

（b） Bilinear

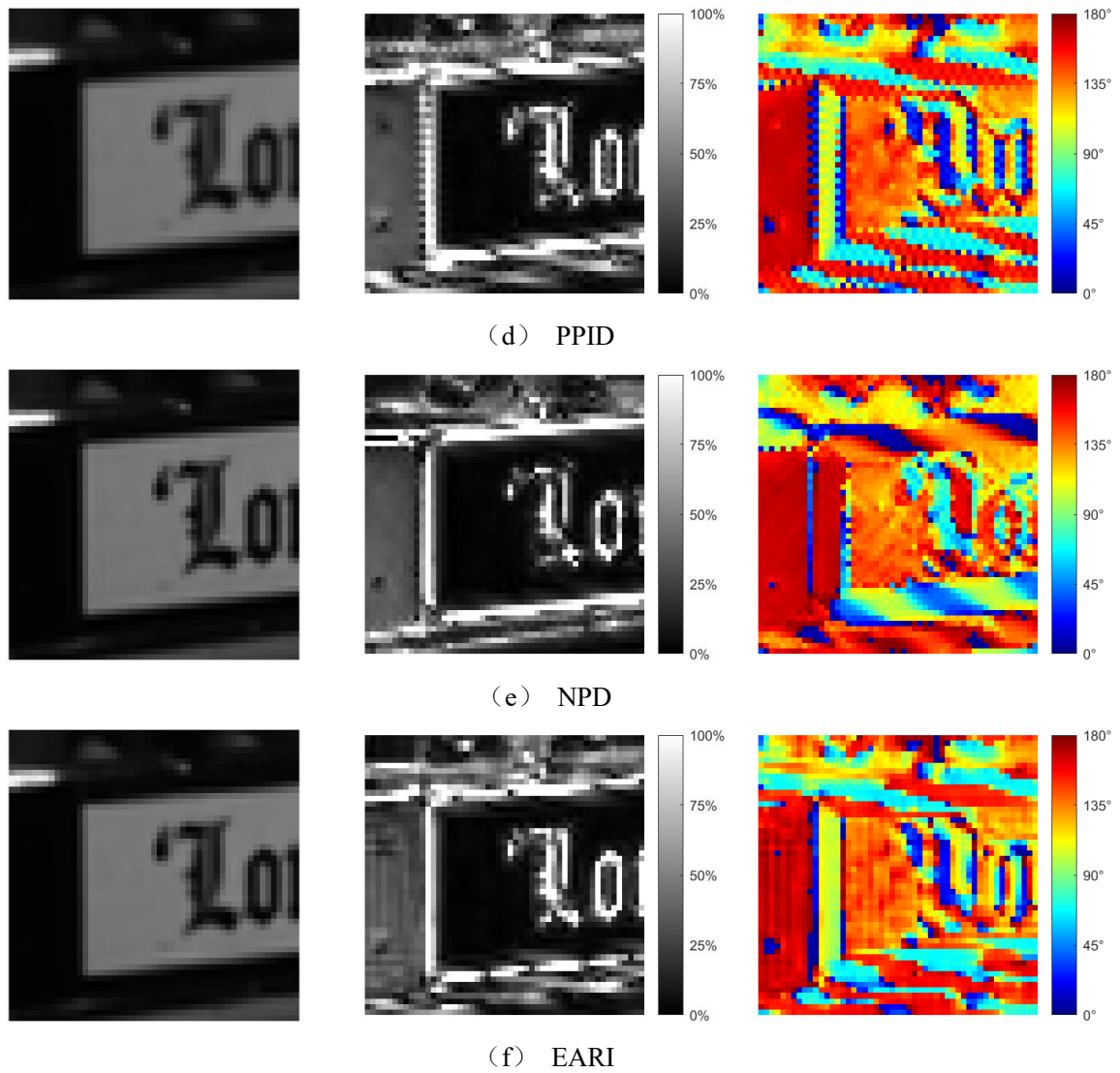（c） ARI

（d） PPID

（e） NPD

（f） EARI

Fig. 4-1 Comparison between the recovered and original images of the same scene by 5 algorithms

In summary, concerning subjective evaluation metrics, all five algorithms perform well in replicating S_0 values, with the Newton interpolation method and the edge aware residual interpolation method being the most effective, especially in edge clarity. In DoLP (degree of linear polarization) replication, all five algorithms produce noticeable pseudo-edges, with the Newton interpolation method having fewer pseudo-edges and achieving closer DoLP value restoration. However, in AoLP restoration, all five algorithms exhibit pronounced pseudo-edges, with the edge aware residual interpolation method demonstrating the best performance. Other algorithms exhibit suboptimal results due to severe pseudo-edges, zipper noise, and inaccuracies in restored values. Therefore, the Newton polynomial interpolation method and the edge knowledge residual interpolation method continue to perform best based on subjective evaluation metrics.

4.3 Summary of this chapter

In this chapter, we employed publicly available image databases and conducted simulation-based experiments using MATLAB to simulate demosaicking algorithms. We utilized PSNR as an objective evaluation metric for image restoration performance. We conducted a comparative analysis of the results from all algorithms, while also referencing summaries and analyses from other literature, as well as considering subjective image evaluation results. Our findings indicate that employing residual-based methods and leveraging inter-channel correlations often leads to favorable results in interpolation. This approach can serve as a promising direction for future algorithmic innovations. Notably, the Newton polynomial interpolation method and the edge-aware interpolation method demonstrated outstanding performance. These algorithms exhibited remarkable results, excelling in both objective metrics and subjective evaluations.

Chapter 5 Summary and Outlook

5.1 Summary of this thesis work

The paper begins by discussing the significance and current development status of acquiring polarized images. It highlights the dominance of DoFP (division of focal plane) technology in polarized image acquisition hardware due to its compact size, integration capabilities, affordability, and suitability for single-shot imaging. DoFP sacrifices resolution to capture more polarization information in a single exposure. It employs a specialized interleaved arrangement of polarization filter arrays, typically with polarization angles at 0°, 45°, 90°, and 135°. Each pixel captures information from only one polarization angle, resulting in incomplete mosaic-like data. Algorithms are then used to restore the complete four-channel polarization information at each point, leveraging properties like inter-channel correlation and spatial relationships, for subsequent computation of key polarization parameters.

The paper primarily focuses on traditional interpolation-based mosaic artifact removal, categorizing them into five main groups: spatially invariant linear filter methods, adaptive weighted interpolation methods, edge-aware interpolation methods, difference-based methods, residual-based methods, and frequency domain replacement methods. In total, 22 polarized image mosaic artifact removal algorithms are covered. Many of these algorithms share underlying principles and processes, thus the paper logically presents algorithm principles, derivations, innovations, strengths, and weaknesses.

Finally, the paper employs MATLAB simulations using publicly available image databases or publicly available code for all the mentioned algorithms. The peak signal-to-noise ratio (PSNR) is used as an objective evaluation metric for comparing algorithm results. The paper compares the outcomes of all algorithms while cross-referencing differences from the results in related survey articles from 2018. Subjective evaluation metrics are also considered to discuss the pros and cons of different algorithms.

The findings reveal that the Newton polynomial interpolation method and the edge-aware residual interpolation method exhibit higher signal-to-noise ratios, making them reliable reference algorithms for polarized image mosaic artifact removal. Moreover, combining difference-based (or residual-based) methods with the notion of inter-channel correlation can lead to improved mosaic artifact removal results, providing valuable guidance for future algorithmic improvements and innovations.

5.2 Outlook

This paper focused exclusively on traditional interpolation-based mosaic artifact removal methods. Currently, some literature utilizes deep learning techniques for mosaic artifact removal, achieving superior results compared to traditional methods. However, these deep learning approaches often come with longer processing times and require a certain level of expertise in both deep learning and programming. Subsequent efforts will concentrate on exploring the application of deep learning in mosaic artifact removal.

Moreover, numerous studies have integrated denoising and mosaic artifact removal, yielding better restoration results compared to standalone approaches. Future work will also encompass research into denoising methods and how to effectively apply them to polarized image mosaic artifact removal.

The selection of objective evaluation metrics in this paper is limited, and it may not comprehensively assess algorithms from all angles. Common evaluation metrics for polarized mosaic artifact removal, in addition to PSNR, include SSIM (structural similarity) and MTF (modulation transfer function). SSIM primarily assesses the similarity between polarized images before and after restoration, while MTF is often used to intuitively reflect the preservation and loss of different frequency information in mosaic artifact removal algorithms.

Subsequent work will introduce a wider variety of evaluation metrics to better assess the strengths and weaknesses of algorithms. Furthermore, the availability of more publicly accessible polarized image databases is anticipated, which will help mitigate algorithm simulation errors resulting from limited sample sizes.

References

Wan C. 2022. Research on infrared polarization image processing algorithm based on FPGA. [Master Thesis for Professional Degree]. Chengdu: University of Electronic Science and Technology of China

Zhang S, Sun Z, Mu Q, Peng Z, Liu H, Liu C. 2022. Review on liquid crystal micropolarizer array for polarization imaging. *Chinese Journal of Liquid Crystals and Displays*, 37（3）: 292~309

Adams J E, Hamilton J F. 1997-05-13. Adaptive color plane interpolation in single sensor color electronic camera. US Utility Patent, 5629734

Ahmed A, Zhao X, Gruev V, Zhang J, Bermak A. 2017. Residual interpolation for division of focal plane polarization image sensors. *Optics Express*, 25（9）: 10651~10662

Brauers J, Aach T. 2006. A color filter array based multispectral camera. *12. Workshop Farbbildverarbeitung*, 55~64

Gao S, Gruev V. 2011. Bilinear and bicubic interpolation methods for division of focal plane polarimeters. *Optics Express,* 19（27）: 26161~26173

Gao S, Gruev V. 2013. Gradient-based interpolation method for division-of-focal-plane polarimeters. *Optics Express*, 21（1）:1137~51

He K, Sun J, Tang X. 2012. Guided image filtering. *IEEE transactions on pattern analysis and machine intelligence*, 35（6）:1397~409

Kiku D, Monno Y, Tanaka M, Okutomi M. 2013. Residual interpolation for color image demosaicking. *2013 IEEE international conference on image processing*, 2304~2308

Kiku D, Monno Y, Tanaka M, Okutomi M. 2016. Beyond color difference: Residual interpolation for color image demosaicking. *IEEE Transactions on Image Processing*, 25（3）: 1288~1300

Lapray P J, Gendre L, Bigué L, Foulonneau A A. 2018. Database of Polarimetric and Multispectral Images in the Visible and NIR Regions. *Unconventional Optical Imaging*, （10677）: 666~679

Lapray P J, Gendre L, Foulonneau A, Bigué L. 2018. An FPGA-based pipeline for micropolarizer array imaging. *International Journal of Circuit Theory and Applications,* 46: 1675~1689

LeMaster D A, Hirakawa K. 2014. Improved microgrid arrangement for integrated imaging polarimeters. *Optics Letters*, 39: 1811~1814

Li N, Zhao Y, Pan Q, Kong S G. 2019. Demosaicking DoFP images using Newton's polynomial interpolation and polarization difference model. *Optics Express*, 27（2）: 1376~1391

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

Miao L, Qi H, Ramanath R, Snyder W E. 2006. Binary tree-based generic demosaicking algorithm for multispectral filter arrays. *IEEE Transactions on Image Processing*, 15（11）: 3550~3558

Mihoubi S, Lapray P J, Bigué L. 2018. Survey of demosaicking methods for polarization filter array images. *Sensors*, 18（11）: 3688

References

Mihoubi S, Losson O, Mathon B, Macaire L. 2015. Multispectral demosaicking using intensity-based spectral correlation. *2015 International Conference on Image Processing Theory, Tools and Applications,* 461~466

Mihoubi S, Losson O, Mathon B, Macaire L. 2017. Multispectral demosaicking using pseudo-panchromatic image. *IEEE Transactions on Computational Imaging,* 3（4）:982~995

Monno Y, Kiku D, Tanaka M, Okutomi M. 2015. Adaptive residual interpolation for color image demosaicking. *2015 IEEE International Conference on Image Processing*, 3861~3865

Morimatsu M, Monno Y, Tanaka M, Okutomi M. 2020. Monochrome and color polarization demosaicking using edge-aware residual interpolation. *2020 IEEE International Conference on Image Processing*, 2571~2575

Nordin G P, Meier J T, Deguzman P C, Jones M. 1999. Diffractive optical element for Stokes vector measurement with a focal plane array. *Polarization Measurement, Analysis, and Remote Sensing II,* 3754: 169~177

Qiu S, Fu Q, Wang C, Heidrich W. 2019. Polarization demosaicking for monochrome and color polarization focal plane arrays.

Ratliff B M, LaCasse C F, Tyo J S. 2009. Interpolation strategies for reducing IFOV artifacts in microgrid polarimeter imagery. *Optics Express*, 17（11）: 9112~9125

Ratliff B M, LaCasse C F, Tyo J S. 2011. Adaptive strategy for demosaicking microgrid polarimeter imagery. *2011 Aerospace Conference*, 1~9

Ratliff B M, Tyo J S, Boger J K, Black W T, Bowers D L, Fetrow M P. 2007. Dead pixel replacement in LWIR microgrid polarimeters. *Optics Express*, 15（12）: 7596~7609

Tyo J S, Goldstein D L, Chenault D B, Shaw J A. 2006. Review of passive imaging polarimetry for remote sensing applications. *Applied Optics,* 45（22）: 5453~5469

Wang X, Thomas J B, Hardeberg J Y, Gouton P. 2013. Discrete wavelet transform based multispectral filter array demosaicking. *2013 Colour and Visual Computing Symposium,* 1~6

Wikipedia. 2023. Bicubic interpolation. https://en.wikipedia.org/wiki/Bicubic_interpolation[2023-3-25]

Wu R, Zhao Y, Li N, Kong S G. 2021. Polarization image demosaicking using polarization channel difference prior. *Optics Express*, 29（14）: 22066~22079

Xin J, Li Z, Wu S, Wang S. 2023. Demosaicking DoFP images using edge compensation method based on correlation. *Optics Express*, 31（9）: 13536~13551

Zhang J, Luo H, Hui B, Chang Z. 2016. Image interpolation for division of focal plane polarimeters with intensity correlation. *Optics Express,* 24（18）: 20799~20807

Zhang J, Ye W, Ahmed A, Qiu Z, Cao Y, Zhao X. 2017. A novel smoothness-based interpolation algorithm for division of focal plane Polarimeters. *2017 IEEE International Symposium on Circuits and Systems*,1~4

Acknowledgements

Time flies by, and in the blink of an eye, the carefree days of my undergraduate years are coming to an end. Though four years may seem short, the invaluable lessons and experiences I've gained at Northwest A&F University have been truly enriching. As graduation approaches, there are many people I would like to express my gratitude to. It is your support and assistance that have given me the motivation to continually progress.

First and foremost, I want to thank my advisor, Associate Professor Wenyi Ren, for his dedicated guidance and high standards. Through his mentorship, I have been able to accumulate knowledge and insights in research and my field of study. From writing and revising my first paper to submitting it for publication, Professor Ren personally guided me through every step, investing his time and effort into each detail. Under his tutelage, I have gradually mastered the norms and requirements of academic writing and developed the ability to conduct thorough research. Professor Ren's profound knowledge, rigorous scholarly attitude, practical work ethic, exemplary character, and unceasing spirit of exploration and innovation will continue to inspire my future endeavors. My heartfelt thanks go to Professor Ren!

I am grateful to all the professors who have taught me over these four years. Each of you has imparted knowledge and explanations with exceptional brilliance, deepening my understanding of optoelectronics and nurturing my passion for further research in this field during my future graduate studies.

I extend my gratitude to my parents, who have nurtured me throughout the years. In my journey of growth, they have consistently set a responsible and caring example, providing unwavering support in all aspects of my life, studies, and finances. Whenever I faced challenges or setbacks, they listened and comforted me, offering consistent encouragement that drives me forward.

Lastly, I appreciate all the professors who have reviewed my papers. Despite their busy schedules, they have generously provided guidance and assistance.

Academic papers published and research achievements during the study period

Lu Y, Jin Z, Ren W. 2023. Pendant drop method to measure surface tension of waterdrops based on smartphone and Matlab at home. *College Physics*, 42（1）: 48~54

Gu J, Ma Y, Li X, Li S, Chen S, Cao Y, Lu Y, Zhang R, Zhou W, Wang H, Jiang J. 2023. Synthesis of Rape Pollen-Fe2O3 Biohybrid Catalyst and Its Application on Photocatalytic Degradation and Antibacterial Properties. *Catalysts*, 13（2）: 358