# A Geometric Model for Polarization Imaging on Projective Cameras

Mara Pistellato[1*] and Filippo Bergamasco[1]

[1*]DAIS, Ca'Foscari University of Venice, 155 via Torino, Venezia, 30030, VE, Italy.

*Corresponding author(s). E-mail(s): mara.pistellato@unive.it;
Contributing authors: filippo.bergamasco@unive.it;

**Abstract**

The vast majority of Shape-from-Polarization (SfP) methods work under the oversimplified assumption of using orthographic cameras. Indeed, it is still unclear how Stokes vector projection behaves when the incoming rays are not orthogonal to the image plane. In this paper, we try to answer this question with a new geometric model describing how a general projective camera captures the light polarization state. Based on the optical properties of a tilted polarizer, our model is implemented as a pre-processing operation acting on raw images, and a scene-independent rotation of the reconstructed normal field. Moreover, our model is consistent with state-of-the-art forward and inverse renderers (as Mitsuba3 and ART), intrinsically enforces physical constraints among the captured channels, and handles the demosaicing of DoFP sensors. Experiments on existing and new datasets demonstrate the accuracy of the model when applied to commercially available polarimetric cameras.

**Keywords:** polarization, projective camera, shape from polarization, Stokes vector

## 1 Introduction

When light reflects on a surface, its polarization changes according to well-established physical rules that describe such interaction. Among all the factors involved in the process, the surface's intrinsic properties and its orientation also determine the final polarization state of the captured beam.

The literature counts several works that exploit such optical principle, aiming at recovering surface properties from images taken with a rotating linear polarizer in front of the camera (see Wolff and Boult (1993); Atkinson and Hancock (2006); Miyazaki et al. (2003)). Recently, the availability of Division-of-Focal-Plane (DoFP) cameras allows to capture the polarization state of the scene with a single shot and raised the interest of the Computer Vision community in polarization-related applications. Some of these consist in using polarization cues to perform inspection in industrial settings as shown in Meriaudeau et al. (2008); Morel and Gorria (2006), while others allow to perform material classification Tominaga and Kimachi (2008); Wolff (1990); Chen and Wolff (1998) or transparent objects analysis as in Miyazaki and Ikeuchi (2007).

Moving to 3D applications, some works propose to use polarimetric data to perform relative pose estimation Cui et al. (2019) and three-view geometry Chen et al. (2018), while some methods propose to extend the dense monocular SLAM

Yang et al. (2018). Finally, some works propose recovering both surface geometry and refractive index via multispectral polarimetric imaging Huynh et al. (2013).

Among these applications, one popular task is the so-called Shape from Polarization (SfP), where light polarization is exploited to recover per-pixel surface normals and proceed with subsequent surface reconstruction. Such methods are designed to recover the 3D shape of acquired objects from a single view thanks to the information intrinsically encoded in the polarimetric images (see Miyazaki et al. (2003); Shakeri et al. (2021); Baek et al. (2018); Taamazyan et al. (2016)). Several approaches in the SfP domain combine polarimetric imaging with other cues coming from classical techniques such as stereo Fukao et al. (2021), multi-view Atkinson and Hancock (2005); Cui et al. (2017); Zhao et al. (2020); Chen et al. (2018), shading and light constraints Ngo Thanh et al. (2015); Smith et al. (2018) or coarse depth maps Kadambi et al. (2017). Moreover, some works presented in the recent past propose data-driven approaches to perform shape from polarization Ba et al. (2020); Lei et al. (2022) exploiting specially-made datasets that involve the use of polarimetric cameras and 3D scanners capturing a wide range of subjects, from small objects to entire buildings.

Despite the undeniable contribution of the mentioned approaches in the present-day relevant literature, almost all of them rely on the quite unrealistic assumption of operating with an orthographic camera. Indeed, the basic equations often presented in such works are designed for a model where light rays hit the sensor perpendicularly: this is pointed out in some early papers such as in Rahmann and Canterakis (2001); Rahmann (2000).

The majority of proposed approaches employ pinhole cameras, in which light beams hit the image plane with an angle that depends on the camera geometry. In practice, assuming an orthographic model while acquiring with a projective camera leads to non-negligible errors and deformations, especially when we are interested in areas near the image borders or if we have short focal lengths. The authors in Chen et al. (2022); Lei et al. (2022) highlight such a problem and try to formulate a solution for the perspective deformation and the representation of polarization data for non-orthographic devices. Finally, (Lu et al., 2019) explicitly considers perspective projection effects but without taking into account the degree of polarization.

## 1.1 Related Works

In the literature we find an abundance of Shape from Polarization methods, but very few works trying to understand how to deal with projective cameras. Indeed, we are concerned that almost all the methods assume (explicitly or even implicitly) to observe an orthographic projection, ignoring how much this assumption might affect the results.

Only recently two works addressed this problem, providing two different solutions. Chen et al. (2022) gave the first geometric relationship between the polarization phase angle and the azimuth angle of the surface normal. Albeit interesting, the model has two limitations. First, it does not account for normal elevation so its applicability is limited to some specific contexts (i.e. single-view recovery of planar surfaces or multi-view estimation of normals). Second, it does not provide a direct description of how the full polarization state (i.e. the Stokes vector) is captured, but only how the Angle of Linear Polarization (AoLP) is affected by the ray direction.

The second work, proposed by Lei et al. (2022), is based on the observation that "the polarization representation is highly influenced by the viewing direction". Their solution consists of a Convolutional Neural Network taking in input the captured image and a viewing encoding providing cues to the camera's intrinsic properties. Accuracy of the resulting normal maps is currently unmatched, but such data-driven approaches give no explicit information on how the model works internally. As often happens with learning-based solutions, the resulting model is a black box that hardly generalises to different contexts.

## 1.2 Our contributions

In this paper we give the first complete mathematical description of how the scene polarization state is modelled within a projective camera. Indeed, the main contribution of this work consists of the formal analysis of what happens when light rays are captured by a pinhole camera, maintaining full compliance with optics theory.

Our model represents light rays as Stokes vectors via Mueller calculus (see Goldstein (2017) for details), and builds on top of the optical properties of the tilted polarizer presented in Korger et al. (2013) to describe how the polarization state is represented and transformed when interacting with the camera optical elements. We address the same problem as Chen et al. (2022), but we effectively model the incoming light as a full Stoke vector instead of just AoLP. Unlike other approaches, the proposed formulation is not based on empirical evidence (i.e. data-driven) but on a solid physical background describing how light is transmitted. One of the key advantages of the proposed model is that it results in an image pre-processing operation and per-pixel fixed transformation to be applied on the estimated normal field. Therefore, it can be embedded in existing SfP methods designed with orthographic assumption. The proposed operations are straightforward as they do not include parameters to be calibrated or some scene-dependent adjustments since they only depend on the camera's intrinsic parameters. We adopted the same conventions assumed in popular direct and inverse renderers, believing that this would simplify the creation of synthetic datasets closely simulating what can be captured with a real polarimetric camera. This topic is of pivotal importance to training learning-based models and testing existing methods against a controllable Ground Truth. Finally, the model is valid for DoFP and Division-of-Time (DoT) cameras, and generalizes to any number of linear polarizers involved in the acquisition.

# 2 Preliminaries

To understand the theoretical background of the proposed model, we briefly summarize some basic notions about light polarization. The goal here is to highlight crucial aspects that are sometimes neglected when approaching shape-from-polarization. Please refer to Collett (2005); Bass et al. (2009) to learn more about these concepts.

## 2.1 Light polarization

Any visible light ray consists of two orthogonal electric field components $E_x, E_y$ oscillating in the plane transverse to the propagation direction $\vec{k}$. Without loss of generality, $\vec{k}$ can be set coincident with the $z$-axis, so that $E_x, E_y, z$ form an orthogonal reference system in which:

$$\begin{aligned} E_x(z,t) &= A_x \cos\left(\omega t - \frac{2\pi}{\lambda} z + \delta_x\right) \\ E_y(z,t) &= A_y \cos\left(\omega t - \frac{2\pi}{\lambda} z + \delta_y\right). \end{aligned} \tag{1}$$

$A_x$ and $A_y$ are two wave amplitudes, $\omega$ is the angular frequency, $\lambda$ is the wavelength and $\delta_x, \delta_y$ are two arbitrary phases. The point $\big(E_x(z,t)\ E_y(z,t)\big)$ traces a so called *polarization ellipse* when discarding the time-space propagator $\omega t - \frac{2\pi}{\lambda} z$ (i.e. when observing the two waves "projected" in the $x-y$ plane). Orientation angle and eccentricity of such ellipse describe the polarization state of light. For example, when $A_y = 0, A_x > 0$ the optical field oscillates horizontally and we have a *linearly horizontal polarized light* (polarization ellipse degenerates to a horizontal segment). When $A_x = A_y, \delta_x - \delta_y = 0$ we have polarized light with an *Angle of Linear Polarization* (AoLP) of 45°.

An important thing has to be noted here. When we discuss about polarization angles (horizontal, 45°, etc.) we have to make clear the orthogonal reference system in which the electric field components are expressed. The third axis is implicitly known since we *always assume the $z$-axis being the direction of propagation*. The $x$-axis can be any unitary vector orthogonal to $z$, thus providing a reference to which such angles are expressed. Since the system is orthogonal, the $y$-axis is uniquely determined as $y = x \times z$. Therefore, a light ray with an AoLP of 45° can be seen as a ray with an AoLP of 0° if we rotate the reference system 45° counter-clockwise around the $z$-axis.

## 2.2 Stokes parameters and Mueller matrices

$E_x, E_y$ are not directly measurable, so a different formulation is usually preferred. Taking a time average of Eqs. 1 yields the definition of the four quantities:

$$\begin{aligned} S_0^{(p)} &= A_x^2 + A_y^2 \\ S_1^{(p)} &= A_x^2 - A_y^2 \end{aligned}$$

$$S_2^{(p)} = 2A_xA_y\cos(\delta_y - \delta_x)$$
$$S_3^{(p)} = 2A_xA_y\sin(\delta_y - \delta_x) \quad (2)$$

representing the *Stokes polarization parameters* of purely monochromatic coherent light radiation.

Describing light properties with a Stokes vector $S$ is powerful because we can represent partially polarized light as a mixture of unpolarized $(S_0\ 0\ 0\ 0)^T$ and completely polarized light:

$$S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = (1-\rho)\begin{pmatrix} S_0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \rho S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} \quad (3)$$

where

$$\rho = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0}, \quad 0 \leq \rho \leq 1 \quad (4)$$

is called *Degree of Linear Polarization* (DoLP)[1]. It follows from Eq.3 that $S_0$ is the intensity of light and the AoLP $\phi$ is given by

$$\phi = \frac{1}{2}\text{atan2}(S_2, S_1). \quad (5)$$

Stokes parameters can be measured by letting the light rays pass through special materials called polarizers and retarders. Such elements transform the input Stoke vector by means of a linear transformation described by a $4 \times 4$ Mueller matrix $\mathbf{M}$. For example, the Mueller matrix of an ideal linear polarizer with *transmitting axis* oriented with an angle $\alpha$ with respect to the $x$-axis is:

$$\mathbf{M}_\alpha = \frac{1}{2}\begin{pmatrix} 1 & \cos 2\alpha & \sin 2\alpha & 0 \\ \cos 2\alpha & \cos^2 2\alpha & \sin 2\alpha \cos 2\alpha & 0 \\ \sin 2\alpha & \sin 2\alpha \cos 2\alpha & \sin^2 2\alpha & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (6)$$

Also in this case, the reference frame matters and cannot be chosen arbitrarily. The Mueller matrix of a polarizing element must be defined in a system coincident with the one in which the input/output Stokes vectors are expressed. If that is not the case, a rotator $R$ (around the $z$-axis) must be applied to align the reference systems, thus getting:

$$S' = R^T\mathbf{M}RS$$

where $S$ and $S'$ are input and output Stokes vectors respectively. In any case, the $z$-axis of $S, S'$ and $\mathbf{M}$ is fixed to the direction in which the ray travels. We stress the fact that Stokes vectors and Muller matrices alone are meaningless without an associated reference frame.

## 2.3 Polarimetric cameras

A polarimetric camera can measure the first 3 components of a Stokes vector[2] using a set of linear polarizers placed in front of a standard sensor. This can be implemented by adding fixed filters directly onto the pixel grid (DoFP cameras), or taking multiple pictures while rotating a linear polarizer in front of the lenses (DoT cameras). In both cases, the polarizers are parallel to the image plane so that their *transmitting axis* is a vector $\vec{t}_\alpha = \begin{pmatrix}\cos\alpha & \sin\alpha & 0\end{pmatrix}^T$ expressed in the camera reference frame.

Regardless ofthe intrinsic properties of the camera (orthographic or projective), for each pixel a set of $\mathcal{I} = \{I_{\alpha_0}, I_{\alpha_1}, \ldots, I_{\alpha_N}\}$ intensities are captured by using linear polarizers with angles $\mathcal{A} = \{\alpha_0, \alpha_1, \ldots \alpha_N\}$ respectively. The task is to recover the incoming Stokes vector $S = (S_0\ S_1\ S_2\ S_3)^T$ from those observations.

## 2.4 Orthographic model

The orthographic camera model is simple to deal with because all the rays are assumed to enter perpendicularly to the image plane. Therefore, we can conveniently set the reference frame of all the rays such that the first two axes follow the horizontal and vertical ordering of the pixels, and the $z$-axis coincides with the camera optical axis (i.e. the direction in which all the rays are propagating). Note that, since the $y$-axis is commonly oriented downward (following the top-down ordering of image pixels) and the $x$-axes rightward, the polarization angles are measured clockwise from the $x$-axis instead of the classical counterclockwise notation. So, care must be taken when using formulas involving polarization angles.

[1] Eq.4 holds since $S_0^2 = S_1^2 + S_2^2 + S_3^2$ for completely polarized light.

[2] Circular polarization is relatively rare in nature (see Cronin and Marshall (2011)) and therefore it is usually not accounted for.

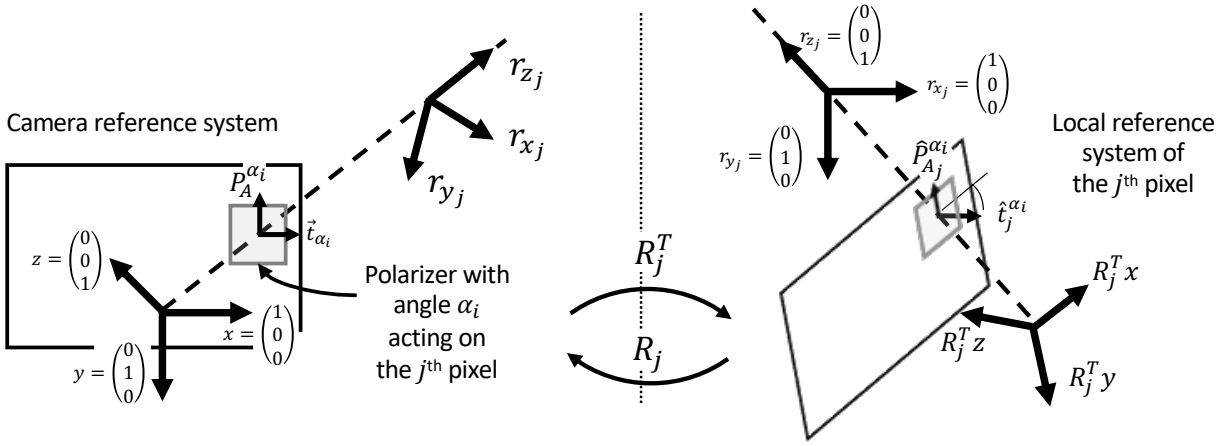


**Fig. 1** Sketch of the proposed model. In the camera reference system (left), the ray passing through the $j^{th}$ pixel (dashed line) is not propagating parallel to the $z$-axis. Therefore, the effect of a polarizer with angle $\alpha_i$ placed parallel to the image plane cannot be described with the orthographic camera model. To solve this, we define a local reference system for each ray (right) oriented along the direction of propagation. Since the linear polarizer is tilted, its effective angle $\hat{\alpha}_i^j$ is different from the actual orientation it has on the sensor. The tilted polarizer model Korger et al. (2013) is used to compensate for this distortion and recover the correct Stokes vector $S_j$ in the reference frame of the $j^{th}$ pixel. Finally, a rotation $R_j$ allows us to map normal vectors estimated from $S_j$ back to the camera reference system.

Since $S_0$ represents the light intensity, for each pixel we can easily relate the intensities $\mathcal{I}$ to the incoming Stokes vector $S$ by writing a set of linear equations:

$$\begin{pmatrix} I_{\alpha_0} \\ I_{\alpha_1} \\ \vdots \\ I_{\alpha_N} \end{pmatrix} = \begin{pmatrix} \mathbf{M}^1_{\alpha_0} \\ \mathbf{M}^1_{\alpha_1} \\ \vdots \\ \mathbf{M}^1_{\alpha_N} \end{pmatrix} \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ 0 \end{pmatrix} \tag{7}$$

where $\mathbf{M}^1_{\alpha_i}$ is the first row of the matrix $\mathbf{M}_{\alpha_i}$ (see Eq. 6). At this point, $S$ can be estimated by solving Eq. 7 in a least-squares sense (see Nayar et al. (1997); Huynh et al. (2010)). This can be done by first expressing each equation $I_{\alpha_i} = \mathbf{M}^1_{\alpha_i} S$ in terms of AoLP $\phi$ and DoLP $\rho$:

$$\begin{aligned} I_{\alpha_i} &= \frac{I_{\max} + I_{\min}}{2} + \frac{I_{\max} - I_{\min}}{2} \cos(2\alpha_i - 2\phi) \\ \rho &= \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}. \end{aligned} \tag{8}$$

A special case is given by Polarimetric Filter Array (PFA) cameras composed by four polarizers arranged with angles $\alpha_{0,1,2,3} = 0°, 45°, 90°, 135°$. In this setting, the trigonometric Eqs. 8 have a simple closed-form solution:

$$\begin{aligned} S_0 &= I_0 + I_{90} \\ S_1 &= I_0 - I_{90} \\ S_2 &= I_{45} - I_{135}. \end{aligned} \tag{9}$$

Note that, due to the orthogonality of such polarizers, intensities are subject to the constraint $I_0 + I_{90} = I_{45} + I_{135}$. So, $S_0$ can alternatively be computed as $I_{45} + I_{135}$.

# 3 The proposed model

The problem of dealing with a projective camera arises because rays propagate in different directions and none of them, except the central one, is parallel to the optical axis. Therefore, we need to:

1. Define a unique reference system for each ray. The $z$-axis must always point to the direction of propagation but we still have freedom of choice for the other two. (Sec. 3.1)
2. Understand what happens to the Stokes vector when a ray traverses a linear polarizer *tilted* with respect to the ray direction (Sec. 3.2)

## 3.1 Local ray reference system

We suppose to know the matrix $\mathbf{K}$ of intrinsic camera parameters that can be estimated with any calibration tool. For each pixel $p_j = \begin{pmatrix} u_j & v_j \end{pmatrix}$ in the image plane, the corresponding *exiting ray* is the

3D unitary vector:

$$\vec{r}_{z_j} = \frac{\mathbf{K}^{-1}\begin{pmatrix} u_j & v_j & 1 \end{pmatrix}^T}{\|\mathbf{K}^{-1}\begin{pmatrix} u_j & v_j & 1 \end{pmatrix}^T\|}. \tag{10}$$

The third axis of $p_j$'s local reference system must be $\vec{r}_{z_j}$ because it represents its direction of propagation[3]. Since the other two can be chosen freely, without loss of generality we set:

$$\vec{r}_{x_j} = \frac{\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^T \times \vec{r}_{z_j}}{\|\begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^T \times \vec{r}_{z_j}\|} \tag{11}$$

$$\vec{r}_{y_j} = \vec{r}_{z_j} \times \vec{r}_{x_j} \tag{12}$$

thus creating the orthogonal reference system $\mathcal{P}_j = (\vec{r}_{x_j}, \vec{r}_{y_j}, \vec{r}_{z_j})$ shown in Fig. 1. Note that there is no physical reason to prefer $\vec{r}_{x_j}$ as defined in Eq. 11 since any other vector orthogonal to $\vec{r}_{z_j}$ would have been equally valid for the model. However, it makes sense to set the frame such that its horizontal axis is aligned with the camera $x$-axis (indeed, $r_{x_j} \perp y$ by construction). Moreover, this is the same convention used by the state-of-the-art polarization aware renderers Mitsuba Jakob et al. (2022) and ARC Wilkie (2018) so synthetically generated images can be easily compared with real images processed with our model. Finally, we define the matrix:

$$R_j = \begin{pmatrix} | & | & | \\ \vec{r}_{x_j} & \vec{r}_{y_j} & \vec{r}_{z_j} \\ | & | & | \end{pmatrix} \tag{13}$$

mapping vectors from $\mathcal{P}_j$ to the camera reference system.

## 3.2 Tilted polarizers

Polarimetric cameras are constructed so that the array of polarizers are parallel to the image plane. Therefore, in each local reference system $\mathcal{P}_j$, such polarizers are tilted and their transmission axis $\vec{t}_{\alpha_i}$ would not be orthogonal to $\vec{r}_{z_j}$. For this reason, Eq. 7 is not correctly defined because each $\mathbf{M}_{\alpha_i}$ lies in the camera reference frame, which is of course not aligned with the ray propagation direction.

[3]To be precise, $-\vec{r}_{z_j}$ is the true direction but changing the sign will not affect the orientation and Mueller calculus still applies.

To solve the problem, we follow the empirical model of Korger et al. (2013) assuming to have polarizing elements made of anisotropic absorbing and scattering particles. According to such model, the *effective transmitting axis* $\hat{\mathbf{t}}_j$ of a tilted polarizer is orthogonal to both $\vec{r}_{z_j}$ and the *absorbing axis* $\hat{P}_{A_j}$ (see Fig. 1). In other words, the effect of a tilted polarizer with angle $\alpha$ is equivalent of a linear polarizer aligned with the ray direction (so that its effect on the Stokes vector can be expressed with a Mueller matrix) but with a different *effective angle* $\hat{\alpha}$. We can obtain the Mueller matrix $\mathbf{M}_{\hat{\alpha}_i^j}$ of the $i^{th}$ tilted polarizer on in the local frame of the $j^{th}$ pixel by computing the effective angle $\hat{\alpha}_i^j$ as follows:

$$\hat{P}_{A_j}^{\alpha_i} = R_j^T \begin{pmatrix} \cos(\alpha_i + \pi/2) \\ \sin(\alpha_i + \pi/2) \\ 0 \end{pmatrix} \tag{14}$$

$$\hat{\mathbf{t}}_j^{\alpha_i} = \begin{pmatrix} \cos\hat{\alpha}_i^j \\ \sin\hat{\alpha}_i^j \\ 0 \end{pmatrix} = \frac{\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T \times \hat{P}_{A_j}^{\alpha_i}}{\|\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T \times \hat{P}_{A_j}^{\alpha_i}\|}. \tag{15}$$

Equation 14 expresses the absorbing axis of the $i^{th}$ polarizer in the local reference frame of the ray. We add $\frac{\pi}{2}$ to the polarizer angle $\alpha_i$ because we assume the absorbing axis being orthogonal to the transmission axis. Then, Eq. 15 computes the effective transmitting axis $\hat{\mathbf{t}}_j^{\alpha_i}$, which is orthogonal to the ray direction of propagation by construction. The angle of $\hat{\mathbf{t}}_j^{\alpha_i}$ in the local reference frame of the ray gives the effective angle $\hat{\alpha}_i^j$.

To summarise, a perspective camera can be used like an orthographic one with the following precautions:

1. Each computed Stokes vector is defined on a different reference frame, depending on the ray direction for that pixel. Consequently, surface normal vectors computed from the Stokes lie on different frames as well, and must be transformed back to the camera reference frame through $R_j$.
2. Even if the camera uses different polarizers with angles in $\mathcal{A}$, each pixel will observe equivalent polarizers with a different set of angles $\hat{\mathcal{A}}_j = \{\hat{\alpha}_1^j, \hat{\alpha}_2^j, \ldots, \hat{\alpha}_N^j\}$. This implies that a linear system like the one shown in Eq. 7 must

be solved in any case, since simpler closed form solutions (See Eq. 9) cannot be valid simultaneously for all the pixels.

### 3.3 How to use our model

The main advantage of our model is that it does not require a reformulation of existing SfP methods designed with the orthographic camera assumption. Indeed, we can synthesize new images that would have been seen with ideal (not tilted) polarizers rotated at angles $0°, 45°, 90°, 135°$. After this, Eq. 9 can be used to get the full Stokes vector and consequently the AoLP and DoLP needed to compute the normal vector field. When normals are recovered, they will be expressed in the local reference frame of each pixel. So, each vector must be transformed to the common camera reference frame by applying the rotation $R_j$ (Eq. 13). We now sketch the basic steps to be performed to embed our model in existing or future approaches for Shape-from-Polarization:

1. Calibrate the camera to get intrinsic matrix $\mathbf{K}$
2. Compute per-pixel reference systems and transformation matrices $R_j$ using Eq. 13 (note that these are fixed and not scene-dependent).
3. For each pixel $j$ and for each polarizer angle $\alpha_i$, compute the effective polarizer angle $\hat{\alpha}_i^j$ using Eq. 15
4. Solve the linear system in Eq. 7 but using $\mathbf{M}^1_{\hat{\alpha}_i^j}$ instead of $\mathbf{M}^1_{\alpha_i}$ to compensate the effect of the tilted polarizers. This will produce a Stokes vector $S_j$ for each pixel. Note that this vector is not expressed in the camera reference frame but in the pixel reference frame $\mathcal{P}_j$.
5. If the SfP method directly accepts the Stokes vector, use the ones computed in the previous step. Otherwise, pre-process the data by synthesizing a new set of images $\hat{\mathcal{I}} = \{\hat{I}_0, \hat{I}_{45}, \hat{I}_{90}, \hat{I}_{135}\}$ where the image $\hat{I}_\gamma$ is obtained by multiplying each Stokes vector $S_j$ with the Muller matrix $\mathbf{M}_\gamma$.
6. Since Stokes vector are in local frames, the 3D normal vectors are expressed in the local reference frames as well. Therefore, whenever a normal $\vec{n}_j$ is estimated for a pixel $j$, it must be transformed back to the camera reference system computing $R_j\vec{n}_j$.

Solving a small linear system for each pixel (Step 4) is the price to pay to compensate the effects of tilted polarizers. However, this operation can be used to seamlessly demosaic a DoFP camera. Indeed, in such cameras, each pixel can only observe a single polarizer angle $\alpha_i$ similar to how a pixel observes a specific color in color cameras with Bayer's pattern. Typically a $2 \times 2$ macro pixels grouping is performed because polarizer angles are conveniently $0, 45, 90$, and $135$ degrees: such angles are however not correct when rays are tilted so it makes sense to correct them before doing any other operation, like computing the Stokes vector. Eq. 7 can be slightly adapted to accept groups of $H \times H$ neighbours of the $j^{th}$ pixel to produce $S_j$. By doing so, we can take into account the tilt of each individual micro-polarizer in the group to obtain the correct Stokes vector, embedding the described correction for tilted polarizers. The approach is similar to Zhang et al. (2016) but can now take into account the effective angles of micro-polarizers. Moreover, the synthesized images (step 5) satisfy physical constraints deriving from the orthogonality of the polarizers. Indeed, it is guaranteed that $\hat{I}_0 + \hat{I}_{90} = \hat{I}_{45} + \hat{I}_{135}$ which is not true in general for raw images in $\mathcal{I}$ regardless the effect of the tilted polarizers. This problem was rarely addressed in the literature but can introduce biases in the computation of Stokes vector since we implicitly give more importance to some polarizer angles (0 and 90 degrees) with respect to the other pair. With our approach this problem completely disappears and corrections (as in Fatima et al. (2022); Pistellato et al. (2023)) are no longer required to obtain coherent values.

## 4 Experiments

In this section we present some experiments to evaluate the ability of the proposed model to accurately describe how Stokes vectors are imaged by a projective camera. The two similar methods available in the literature are the well-known Orthographic model and the recent Perspective Phase Angle (PPA) proposed in Chen et al. (2022).
As discussed before, PPA just relates the light AoLP $\phi$ to the surface normal without giving a unified description of what happens to the whole Stokes vector. Therefore, comparisons against PPA are limited to specific experiments in Sections 4.1 and 4.3. To reduce uncertainty due to uncontrollable scene conditions (mixed

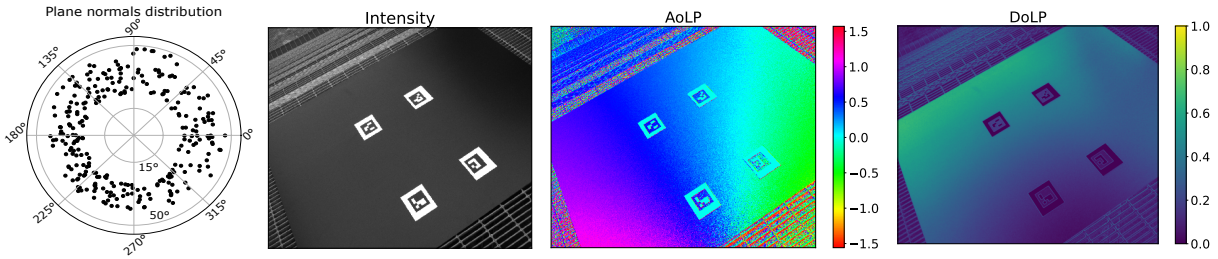


**Fig. 2** From left to right: angular distribution of plane normals in our dataset, an example of intensity image from our dataset with the corresponding AoLP and DoLP.

| | | Model | | |
|---|---|---|---|---|
| | | Our | PPA | Orthographic |
| Our dataset | MAE | **1.88** | 2.71 | 18.04 |
| | RMSE | **2.51** | 3.48 | 21.72 |
| | Median | **1.49** | 2.23 | 16.25 |
| | Accuracy | **0.99 1.00 1.00** | **0.99 1.00 1.00** | 0.65 0.84 0.91 |
| PPA dataset | MAE | **1.54** | 1.95 | 13.60 |
| | RMSE | **2.10** | 2.52 | 16.13 |
| | Median | **1.10** | 1.52 | 11.45 |
| | Accuracy | **0.99 0.99 0.99** | **0.99 0.99 0.99** | 0.76 0.89 0.94 |
| SfP dataset | MAE | **8.63** | 8.77 | 9.52 |
| | RMSE | **13.51** | 13.98 | 14.73 |
| | Median | **5.05** | 5.16 | 5.91 |
| | Accuracy | **0.87 0.94 0.96** | 0.84 0.93 0.94 | 0.83 0.92 0.95 |

**Table 1** AoLP accuracy. We report errors (in degrees) of the expected AoLP computed given the normal vectors of the datasets with respect to the observed AoLP. We tested our perspective model, PPA (Chen et al. (2022)) and Orthographic on three different datasets, namely: our acquired dataset, PPA dataset and deep SfP dataset (see Ba et al. (2020)). The three accuracy values refers to % of pixels with an error under 11.25°, 22.5° and 30°.

polarization, accuracy of the ground truth surface normals, $\frac{\pi}{2}$-ambiguity, etc.) we followed the approach of Chen et al. (2022).

We acquired our dataset using a glossy planar plastic board with markers attached to it so that the plane orientation (i.e. its normal in the camera reference frame) can be accurately recovered. One example of acquired image is shown in Figure 2. We opted for such a setup for several reasons: despite being simple, a plane is ideal to evaluate the correctness of a model, as the surface normal is constant and the effect of perspective projection becomes easily detectable. Moreover, working with complex objects and 3D scanners would have introduced several sources of noise in our analysis, that are not relevant to support our claims. Thus, we created a new dataset similar to PPA dataset in Chen et al. (2022) using a different camera and lenses. Specifically, we used a FLIR Blackfly camera mounting a 5 Mpixel Sony IMX250MZR DoFP sensor with 8mm lenses (field-of-view $\approx 107°$). Intrinsics camera parameters were calibrated in a standard way as described in Zhang (2000), acquiring a planar checkerboard. The obtained calibration RMS error was $\approx 0.5$ px. The scene was illuminated with natural light on a very overcast day so that the DoLP of incoming light is close to 0 (i.e. incoming light rays are unpolarized). This detail is fundamental for our experimental setup: indeed, when incoming light is polarized, the resulting Stokes vector after reflection is altered according to Fresnel's equations. The sky on a sunny day is far from being unpolarized, as discussed in several works (see Goldstein (2017), Strutt (1871)) and also exploited in Ichikawa et al. (2021). The resulting dataset is composed of 300 images of the plane taken at different angles and distances (See Fig.2, Right). Since our goal was to cover all camera rays, we ensured that all camera pixels captured

the surface multiple times and from different viewing points. Specular reflection for the material is dominant in our data (as in PPA), so the $\pi$-ambiguity is the only one that can happen on the AoLP. Therefore, we took the best between $\phi_m$ and $\phi_m + \pi$.

## 4.1 AoLP accuracy

We started our evaluation by computing the AoLP accuracy. We did that by measuring the agreement between the AoLP $\phi_c$ captured by the camera and the expected AoLP $\phi_m$ computed by the three models (Ortho, PPA, and Our) given the surface normals. In the Orthographic model $\phi_m = \text{atan2}(n_y, n_x)$ (i.e. the AoLP is equal to the azimuth of the surface normal $\vec{n} = (n_x, n_y, n_z)^T$), for PPA is given by Eq. 6 in Chen et al. (2022), and in our model is computed as the Orthographic model but adding the tilted polarizer correction and applying the per-pixel vector rotation as explained in Sec. 3.3. Results are listed in Tab. 1, where we tested the model on three different datasets: our acquired dataset, the dataset proposed in Chen et al. (2022) (PPA dataset) and the one presented in Ba et al. (2020) (SfP dataset). Both PPA and our model return a significantly lower error than the Orthographic model. This result, consistent to what is declared in Chen et al. (2022), should raise the awareness that ignoring perspective distortion is probably not reasonable for accurate Shape-from-polarization. Our model performs better for all datasets because it compensates the effect of tilted polarizers: both MAE and RMSE of our model show an improvement of $\approx$ 25% on PPA dataset and of $\approx$ 40% on our data. The difference between our and PPA models is mainly because the plane in PPA data does not cover the outermost pixels of the image, where the error is more significant (see Fig. 3 for more details). Regarding the comparison on SfP dataset, we need to include some considerations: (i) the camera estimated focal length for this dataset is $\sim$2500 px, that makes the rays less tilted with respect to our setup (that is $\sim$1210 px) and closer to orthographic, (ii) the captured objects are always in the central area of the image, again reducing the available rays angles, (iii) the average signal DoLP is very low, and (iv) polarized light sources are not excluded (as already discussed, when incoming light is polarized, the reflected Stokes is altered). These factors make the dataset not suitable for our purposes: since we are proposing a model describing how Stokes vectors are acquired in a perspective camera, we should assess its validity by excluding external factors and analysing a controlled scene as we did in our data. Nevertheless, the results on SfP dataset are still interesting, because our model exhibits a better angle accuracy with respect to orthographic and PPA models. This result is an additional indication that applying a correct perspective model for polarimetric data is far from being negligible for any application.

In Fig. 3 we show the per-pixel Mean Absolute Error (MAE) for the AoLP (estimated with the three methods) obtained by putting together all the images from our acquired dataset (first row) and PPA dataset (second row). Excluding some artifacts due to highlights that saturated some pixels, we observe a strong radial pattern in both the Orthographic and PPA images (left and center). PPA model performs better than Ortho, but the radial pattern is still visible, especially in our dataset where we managed to cover all the camera pixels, including the external areas of the image. This is expected since rays corresponding to pixels farther away from the principal point are more tilted, and in our model this effect is strongly attenuated, showing evidence that the perspective distortion has been effectively compensated.

In addition, in Figure 4 we show the distance from the image principal point against the average error on the estimated AoLP for our proposed model and PPA. The leftmost plot refers to PPA dataset and the rightmost to our dataset. For both datasets the plots show that when camera rays are significantly tilted (i.e. peripheral image areas), our model offers the best AoLP estimation. Indeed, our perspective model is always under 3 degrees, while PPA reaches errors of 5 degrees in regions far from the central area.

## 4.2 DoLP accuracy

Following the AoLP accuracy experiments, we performed similar tests on the measured DoLP to assess the performance of our model. In this case, we need to compare the expected DoLP (computed from the surface normal) with the DoLP captured by the camera. The operation is not as immediate as in the case of the AoLP, since the

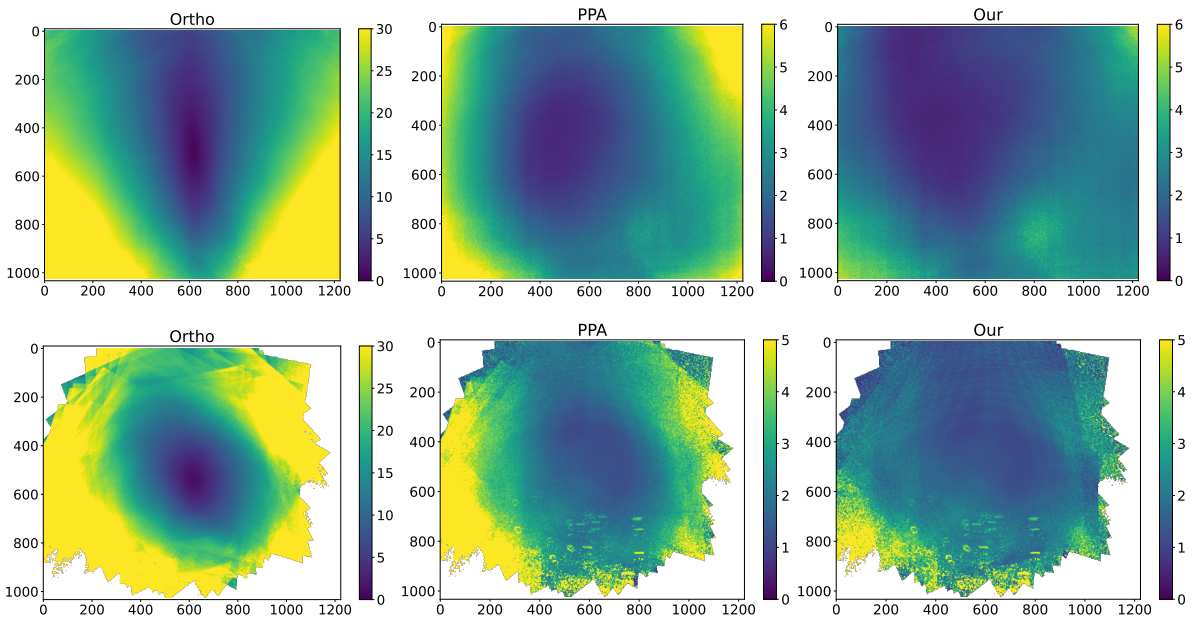


**Fig. 3** Average per-pixel error (in degrees) of the predicted AoLP for three models, namely: orthographic, PPA (Chen et al. (2022)) and ours. The first row shows errors on our dataset, while the second on the dataset presented in Chen et al. (2022). Note that the colour bar for the Orthographic model is substantially different to avoid saturation.

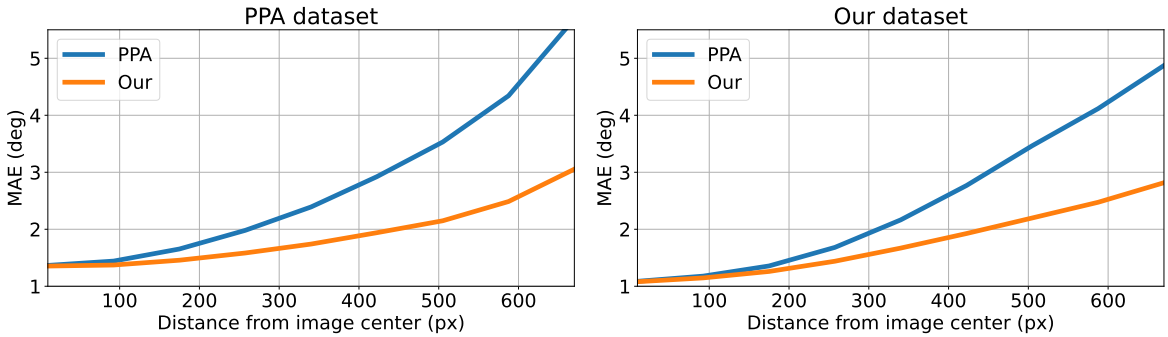


**Fig. 4** AoLP estimation MAE vs. distance from the center for our dataset (left) and PPA dataset (right). The curves show the average error of estimated AoLP with respect to the distance from the central pixel.

DoLP value is related to the zenith angle of the surface normal in a non-linear way.

In our case the specular reflection is dominant, so we used the function shown in Atkinson and Hancock (2006) to relate the DoLP $\rho$ with zenith angle $\theta$ and the index of refraction $n$:

$$\rho(n,\theta) = a\frac{2\sin^2\theta\cos\theta\sqrt{n^2-\sin^2\theta}}{n^2-\sin^2\theta-n^2\sin^2\theta+2\sin^4\theta} \quad (16)$$

where $a$ is a scale parameter we added to the original formulation to account for a possible diffuse component of reflection and other nonlinear contribution like the Umov's effect Kupinski et al. (2019); Umow (1905). To apply Eq. 16 we need first to determine the unknown parameters $n$ and $a$. We estimated such parameters with a data-driven approach by fitting Eq. 16 to the measured DoLP. This is possible under the assumption that: (i) the observed geometry is indeed a plane with a known pose, and (ii) the plane is composed of a material with a uniform index of refraction everywhere[4]. Figure 5 shows the relationship of the zenith angle $\theta$ (on x-axis) of the surface normal in

[4]Since our dataset uses a different plane material than the one used for PPA, we estimated a different set of parameters for each dataset.

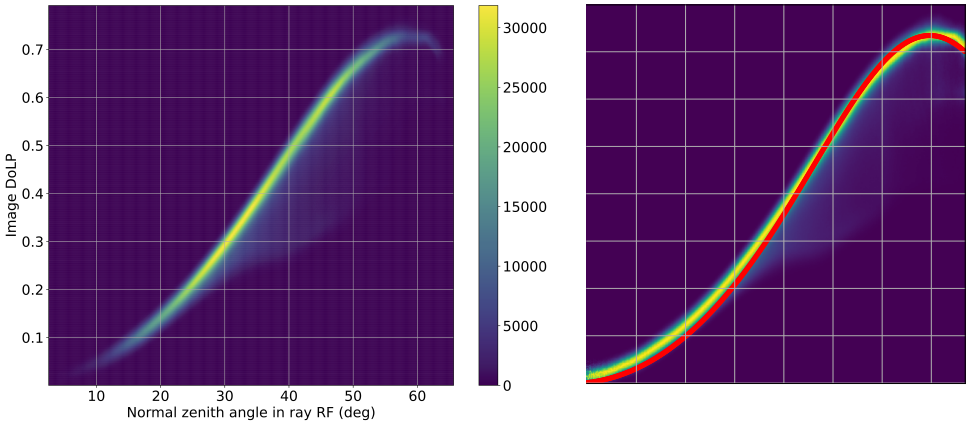


**Fig. 5** Left: 2D histogram computed on our dataset displaying the relationship between the observed zenith angle of the plane normal (in ray reference frame) with the observed DoLP. Right: the same normalized data and the fitted curve given by Eq. 16 optimizing the parameters $a$ and $n$. Moreover, the plot shows that diffuse reflection is negligible for the plastic material under study as the measured DoLP values fit very well with the specular model (thick red curve in the right image).

| | | Model | |
|---|---|---|---|
| | | Our | Orthographic |
| Our dataset | MAE | **0.0350** | 0.1520 |
| | RMSE | **0.0591** | 0.1857 |
| | Median | **0.0201** | 0.1350 |
| PPA dataset | MAE | **0.0167** | 0.0981 |
| | RMSE | **0.0225** | 0.1211 |
| | Median | **0.0129** | 0.0862 |

**Table 2** DoLP accuracy: errors of the expected DoLP computed given the datasets normal with respect to the observed DoLP. We tested our perspective model and orthographic, since PPA (Chen et al. (2022)) does not model the DoLP.

ray reference frame with respect to the measured DoLP in the scene (y-axis) for our dataset. The density image (i.e. 2D histogram) on the left was produced by merging all valid pixels from all the images in our dataset. A density plot has been chosen to avoid noise and outliers that would appear on the scattered data, allowing a better estimation of the DoLP curve. The rightmost plot shows the same data normalized by dividing by the maximum value of each column so that each sample on x axis has the same weight. We then optimized an energy function to maximize the overlap between the function defined by Eq. 16 and the underlying density: the final fitted curve is displayed in red. We obtained $a = 0.73$ and $n = 1.57$ for our dataset and $a = 0.46$ and $n = 1.62$ for PPA dataset.

After fitting the correct DoLP model, we applied it to the data to obtain the expected DoLP and compute the average errors, as we did for the AoLP. Results are listed in Table 2, where we compare our model with just the orthographic one since PPA does not model the DoLP. Results show that in both datasets our perspective model exhibits a significantly higher precision in accordance with the AoLP results. Figure 6 shows the per-pixel average error for both our and orthographic model in the two datasets. Also in this case we observe a significant improvement with our model for both datasets. The error in our dataset is consistently smaller than 0.05, except for the upper-right corner of the image where we observe higher errors due to unexpected reflections on the surface. Errors for the orthographic model are higher and exhibit a radial pattern similar to what we observe for the AoLP.

## 4.3 Plane orientation estimation

When we observe a planar surface, the AoLP is sufficient to recover the plane orientation $\vec{n}$. Let $\phi$ be the azimuth angle of $\vec{n}$. When using the orthographic model, the linear constraint

$$\begin{pmatrix}\sin\phi & \cos\phi & 0\end{pmatrix} \cdot \vec{n} = 0 \tag{17}$$

is commonly used for photo-polarimetric stereo approaches or iso-depth contour tracing. If we have at least $K > 3$ pixels observing the same plane, we can apply our model and exploit such constraint to recover the plane normal $\vec{n}$ (in camera reference frame) from the (corrected) $\phi_j$ observed at each pixel. A solution can be obtained

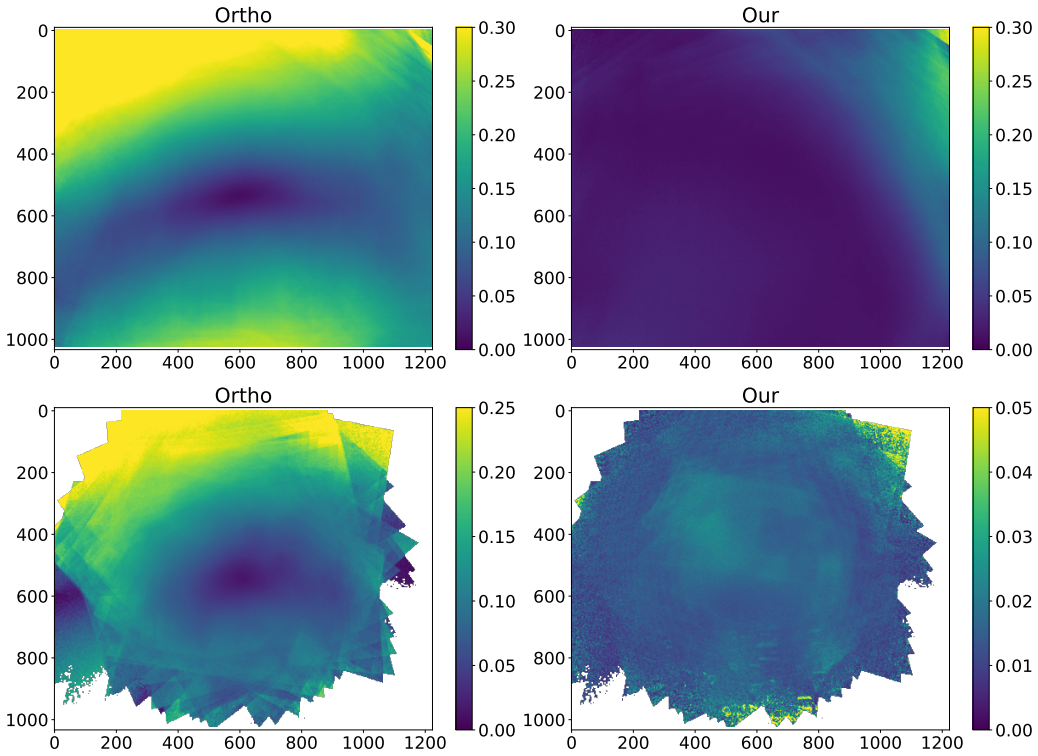


**Fig. 6** Average per-pixel error of the predicted DoLP for orthographic and our model. The first row shows errors on our dataset, while the second on the PPA dataset (Chen et al. (2022)). Note that the colour bar for the Orthographic model is substantially different to avoid saturation.

by solving:

$$\begin{pmatrix} \begin{pmatrix} \sin\phi_1 & \cos\phi_1 & 0 \end{pmatrix} R_1^T \\ \begin{pmatrix} \sin\phi_2 & \cos\phi_2 & 0 \end{pmatrix} R_2^T \\ \vdots \\ \begin{pmatrix} \sin\phi_K & \cos\phi_K & 0 \end{pmatrix} R_K^T \end{pmatrix} \vec{n} = \vec{0} \qquad (18)$$

as a Linear Least-Squares problem where $R_j$ are computed as in Eq. 13 and $\phi_j$ are the corrected AoLPs. A similar constraint is provided for PPA model (see Eq. 7 in Chen et al. (2022)) and solved in the same way. This idea cannot be exploited with the orthographic model because the resulting system would always be under-determined. Indeed, since all rays are parallel, the rotation matrices $R_1, \dots R_K$ will be identities, and each pixel would observe the same $\phi$. In Fig. 7 we plotted the estimated plane normal error distribution against the ground truth data for PPA dataset (Ortho model is not present for the reasons discussed before). Also in this case, our mean absolute error is lower (1.57° vs. 2.89°) and with less variability. This reflects a better estimation of the AoLP due to the correction applied by the tilted polarizer model.

## 4.4 Normal estimation

Our work describes the relation between the incoming Stokes vector and the Stokes vector captured by the camera: this verifies regardless of

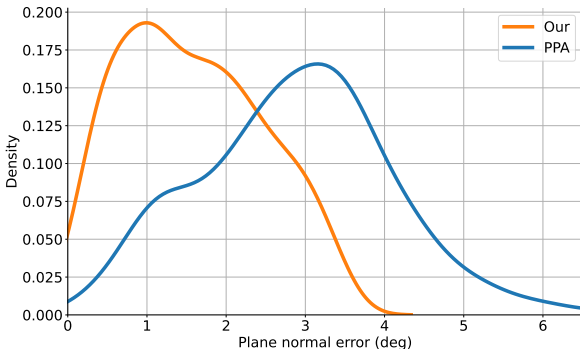


**Fig. 7** Plane normal estimation error distributions for our perspective model and PPA model computed on PPA dataset.

| | PPA Dataset | | Our Dataset | |
|---|---|---|---|---|
| | MAE | RMSE | MAE | RMSE |
| Our | $1.923 \pm 1.343$ | 2.346 | $3.446 \pm 2.638$ | 4.340 |
| Orthographic | $13.441 \pm 6.653$ | 14.997 | $18.423 \pm 8.667$ | 20.35 |
| Smith | $28.494 \pm 12.379$ | 31.067 | $34.848 \pm 5.477$ | 35.275 |
| Smith corrected | $27.396 \pm 12.489$ | 31.021 | $33.870 \pm 6.399$ | 34.470 |
| Mahmoud et al. | $48.19 \pm 10.69$ | 49.37 | $50.058 \pm 8.875$ | 50.839 |
| SfPW | $38.305 \pm 22.915$ | 44.637 | $22.819 \pm 15.649$ | 27.670 |
| DeepSFP | $35.243 \pm 10.758$ | 38.753 | $30.435 \pm 9.81$ | 33.34 |
| DeepSFP corrected | $27.967 \pm 8.880$ | 31.477 | $21.794 \pm 4.439$ | 21.70 |

**Table 3** Comparison of normal estimation errors on PPA and our plane datasets between different SfP methods. Projective (Our) and Orthographic models use an oracle to solve all the ambiguities. Since DeepSfP method uses the orthographic model to compute normal priors, we also present a corrected version where we keep the same CNN weights but computed the normal priors according to our model. Errors are in degrees.

the phenomena that originated the radiation (diffuse/specular reflection, refraction, etc.). Instead, solving the $\pi$-ambiguity to recover the surface normal is a task handled by SfP methods. Since we are not proposing a SfP method, handling such ambiguities is not our goal and thus evaluating our model effectiveness based on how well a normal field is reconstructed has to be done with care.

We recall that the azimuth angle of $\vec{n}$ is related to the AoLP $\phi$ up to a (unavoidable) $\pi$-ambiguity, and an additional $\frac{\pi}{2}$-ambiguity depending if specular or diffuse reflection dominates.

To show the validity of our model, the idea is to test how well it allows the recovery of surface normals in ideal conditions, i.e. with an oracle (a "perfect" SfP approach) that removes ambiguities and provides a close estimate of the parameters $n, a$ in Eq. 16. We estimated the best values for $n, a$ as discussed in the previous DoLP Section, and computed the per-pixel normal vectors from the Stokes vector following the process described in Sec. 3.3 and excluding all the ambiguities.

We compared the resulting normal field against the *orthographic* model using the same oracle as our method (i.e. estimated function $\rho$ and optimal disambiguator), Smith et al. (2018) as in the original paper, Mahmoud et al. (2012), deep SfP from Ba et al. (2020), and SfPW from Lei et al. (2022), a deep CNN which embeds a viewing encoding "to account for non-orthographic projection in scene-level".

Additionally, we propose two modified versions of the methods: in *Smith corrected* we corrected the input image for tilted polarizers and applied the proposed per-pixel rotations to the output vectors, while in *DeepSfP corrected* we also corrected the input image for tilted polarizers and applied the per-pixel rotation to all the normal priors that are given as input to the CNN (with the same weights of the plain DeepSfP). The DeepSfP CNN outputs a normal field taking as input the 4 channels from the PFA camera and three different alternatives for the normals: one obtained applying the specular model and two for each possible diffuse solution. Note that these input vectors are computed with the orthographic model, as described in Ba et al. (2020), and therefore it makes sense to apply our model (i.e. rotate all the normal priors) directly in the network input. We trained the original implementation of DeepSfP on their dataset, and for SfPW we used the original pre-trained model. We do not provide a "corrected" SfPW since the method should already take into consideration perspective effects by encoding the pixel information in the input. Note also that in *Smith corrected* we did not alter the original method, but applied our model afterwards. This of course is sub-optimal because the per-pixel rotations $R_j$ (that are fixed, thus not unknowns to be solved) are applied only at the end.

We listed MAE and RMSE (in degrees) of the estimated normals in Tab. 3: normals recovered with our projective model are significantly more accurate than the others. This is not surprising, since we assume to solve all the unknowns we usually face when doing SfP. The first thing to highlight is the improvement against the orthographic model that uses the same oracle. Indeed, we put the orthographic model in the

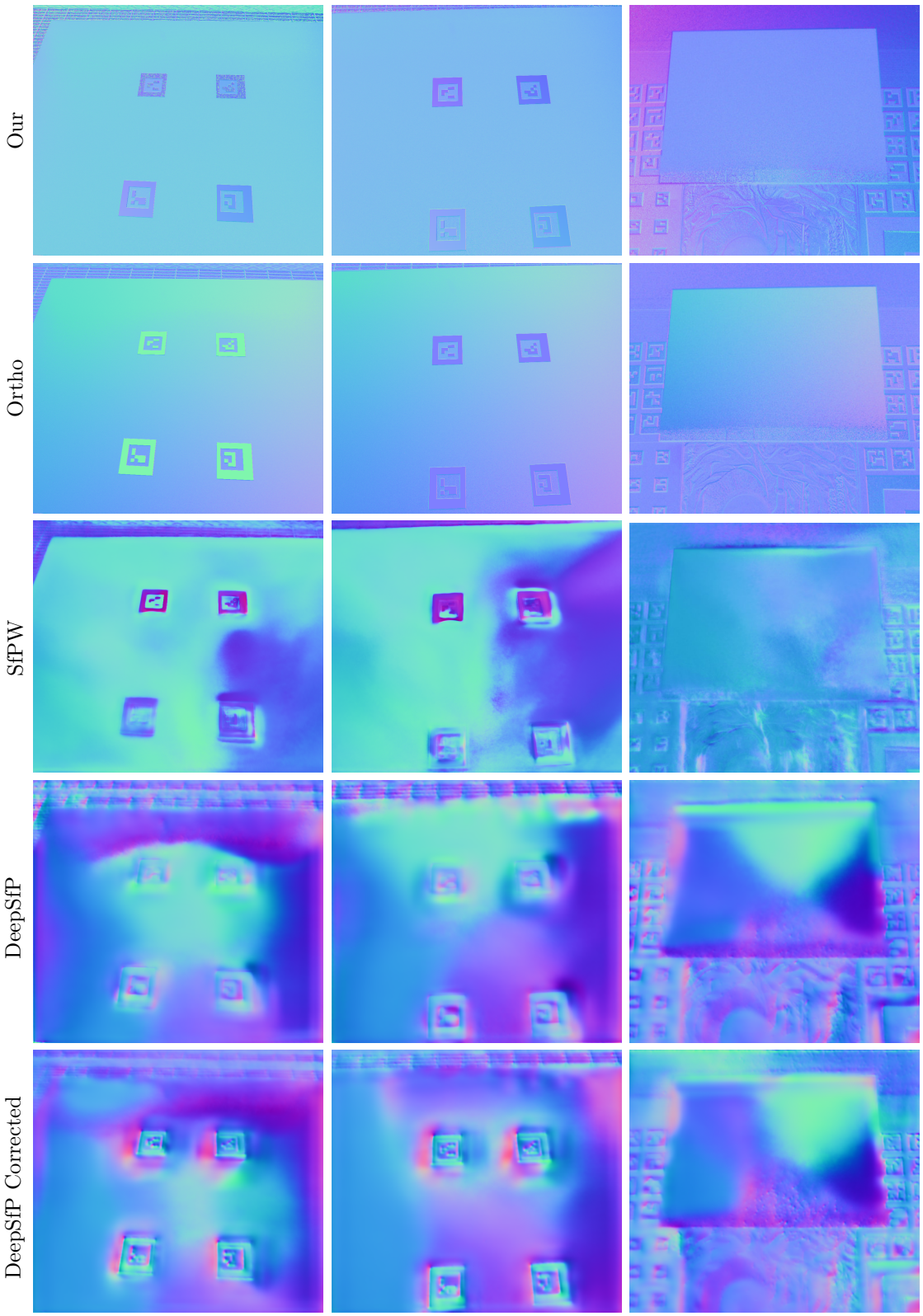


**Fig. 8** Normal maps computed on our and PPA datasets for our model, orthographic model, and data-driven approaches. Errors in the paper are computed inside the plane but we kept the rest of the scene for visualization. We can notice a wrong radial pattern for orthographic model. Learning-based approaches completely fail on this simple planar scene.

same ideal conditions of ours and obtained more than 13° MAE with respect to our perspective model (1.9°). This means that properly accounting for non-orthographic projection is crucial to reduce the final error, no matter how sophisticated the method adopted to solve the ambiguities. Moreover, *Smith corrected* performs better than the original one because accounts for non-orthographic cameras but still suffers from disambiguation errors. When applying DeepSfP to both datasets, the normal estimation errors are higher with respect to other methods. Nevertheless, one very interesting thing that we highlight is the improvement that we obtained from the DeepSfP CNN when we simply correct the input by applying the perspective model. By doing that, we improved the MAE and RMSE of the CNN by almost 10 degrees for both datasets with no particular effort. This is indeed a significant result that shows that the proposed model is effective in representing Stokes vectors for perspective cameras. Finally, SfPW does not perform as well as the "simple" Orthographic model. Considering that SfPW takes into account non-orthographic cameras, we expected a lower error in this experiment. Since SfPW is designed to output a normal map "in the wild", the result is surprisingly bad when tested on a scene with a simple plane. In general, shape from polarization is difficult to solve without posing additional strong priors to the reconstructed scene, and learning-based methods can only try to resolve the ambiguities based on what is observed on the training set, with the risk of overfitting specific 3D structures, and this is clear from the results we obtained for both data-driven approaches.

In Figure 8 we show some qualitative examples of the obtained normal fields for PPA dataset and our acquired data. The captured object is a plane for both datasets, so the normals should be constant on the whole surface. Thanks to the plane setup we can appreciate the radial errors that are clearly introduced in the orthographic model, while our perspective model produces a constant value on the whole surface, especially on peripheral image areas. Regarding the data-driven approaches, SfPW fails by producing non-coherent values on the acquired region, even if it is designed to take into account the camera rays' directions as direct input. Indeed, such behaviour indicates probable data overfitting. The DeepSfP network is shown as the original version in Ba et al. (2020) and with corrected prior input. As we already noted in the comparison table, the corrected input produces a different output, resulting in a smoother normal map for the plane surface, but still, the error is significant.

# 5 Conclusions

In this paper, we presented a model to describe how a projective camera captures the light polarization state. Differently than the empirical PPA model Chen et al. (2022) or learning-based solutions, our formulation directly derives from the optical properties of the tilted polarizer and applies Mueller algebra to model the behaviour of incoming light as a Stoke vector. We consider it a unifying model, equally valid for DoFP (Division of Focal Plane) and DoT cameras (Division of Time, i.e. with rotating polarizers above the sensor or in front of the lenses) and consistent with conventions used by polarization-aware renderers. It allows to embedding demosaicing of DoFP cameras directly in the Stokes estimation process and can be implemented by means of a pre-processing of raw data and rotation of the surface normals in a scene-independent way. In the experimental section, we observed a good agreement between the captured Stokes vector and the expected one in terms of AoLP and DoLP linked to surface normals.

Moreover, the computed normal maps using a perfect disambiguation are coherent with the geometry of the acquired plane, and applying our model to other SfP techniques (as a simple pre- or post-processing) leads to a significant error reduction. In the future, we aim to generalize our model to include lens distortion and vignetting on acquired DoLP due to Umov's effect. This, in turn, would allow a better understanding of how the intrinsic camera parameters affect the measurement of light polarisation state.

**Supplementary information.** The datasets generated during and/or analysed during the current study are available from the corresponding author upon reasonable request.

**Acknowledgments.** This work was partially supported by DAIS - Ca' Foscari University of Venice within the IRIDE program.

# References


Wolff, L.B., Boult, T.E.: Constraining object features using a polarization reflectance model. Phys. Based Vis. Princ. Pract. Radiom **1**, 167 (1993)

Atkinson, G.A., Hancock, E.R.: Recovery of surface orientation from diffuse polarization. IEEE transactions on image processing **15**(6), 1653–1664 (2006)

Miyazaki, D., Tan, R.T., Hara, K., Ikeuchi, K.: Polarization-based inverse rendering from a single view. In: Computer Vision, IEEE International Conference On, vol. 3, pp. 982–982 (2003). IEEE Computer Society

Meriaudeau, F., Ferraton, M., Stolz, C., Morel, O., Bigué, L.: Polarization imaging for industrial inspection. In: Image Processing: Machine Vision Applications, vol. 6813, pp. 72–81 (2008). SPIE

Morel, O., Gorria, P.: Polarization imaging for 3d inspection of highly reflective metallic objects. Optics and Spectroscopy **101**, 11–17 (2006)

Tominaga, S., Kimachi, A.: Polarization imaging for material classification. Optical Engineering **47**(12), 123201 (2008)

Wolff, L.B.: Polarization-based material classification from specular reflection. IEEE transactions on pattern analysis and machine intelligence **12**(11), 1059–1071 (1990)

Chen, H., Wolff, L.B.: Polarization phase-based method for material classification in computer vision. International Journal of Computer Vision **28**(1), 73–83 (1998)

Miyazaki, D., Ikeuchi, K.: Shape estimation of transparent objects by using inverse polarization ray tracing. IEEE Transactions on Pattern Analysis and Machine Intelligence **29**(11), 2018–2030 (2007)

Cui, Z., Larsson, V., Pollefeys, M.: Polarimetric relative pose estimation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 2671–2680 (2019)

Chen, L., Zheng, Y., Subpa-Asa, A., Sato, I.: Polarimetric three-view geometry. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 20–36 (2018)

Yang, L., Tan, F., Li, A., Cui, Z., Furukawa, Y., Tan, P.: Polarimetric dense monocular slam. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3857–3866 (2018)

Huynh, C.P., Robles-Kelly, A., Hancock, E.R.: Shape and refractive index from single-view spectro-polarimetric images. International journal of computer vision **101**(1), 64–94 (2013)

Shakeri, M., Loo, S.Y., Zhang, H., Hu, K.: Polarimetric monocular dense mapping using relative deep depth prior. IEEE Robotics and Automation Letters **6**(3), 4512–4519 (2021)

Baek, S.-H., Jeon, D.S., Tong, X., Kim, M.H.: Simultaneous acquisition of polarimetric svbrdf and normals. ACM Trans. Graph. **37**(6), 268–1 (2018)

Taamazyan, V., Kadambi, A., Raskar, R.: Shape from mixed polarization. arXiv preprint arXiv:1605.02066 (2016)

Fukao, Y., Kawahara, R., Nobuhara, S., Nishino, K.: Polarimetric normal stereo. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 682–690 (2021)

Atkinson, G.A., Hancock, E.R.: Multi-view surface reconstruction using polarization. In: ICCV, vol. 2, p. 3 (2005)

Cui, Z., Gu, J., Shi, B., Tan, P., Kautz, J.: Polarimetric multi-view stereo. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1558–1567 (2017)

Zhao, J., Monno, Y., Okutomi, M.: Polarimetric multi-view inverse rendering. In: European Conference on Computer Vision, pp. 85–102 (2020). Springer

Ngo Thanh, T., Nagahara, H., Taniguchi, R.-i.: Shape and light directions from shading

and polarization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2310–2318 (2015)

Smith, W.A., Ramamoorthi, R., Tozza, S.: Height-from-polarisation with unknown lighting or albedo. IEEE transactions on pattern analysis and machine intelligence **41**(12), 2875–2888 (2018)

Kadambi, A., Taamazyan, V., Shi, B., Raskar, R.: Depth sensing using geometrically constrained polarization normals. International Journal of Computer Vision **125**(1-3), 34–51 (2017)

Ba, Y., Gilbert, A., Wang, F., Yang, J., Chen, R., Wang, Y., Yan, L., Shi, B., Kadambi, A.: Deep shape from polarization. In: European Conference on Computer Vision, pp. 554–571 (2020). Springer

Lei, C., Qi, C., Xie, J., Fan, N., Koltun, V., Chen, Q.: Shape from polarization for complex scenes in the wild. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12632–12641 (2022)

Rahmann, S., Canterakis, N.: Reconstruction of specular surfaces using polarization imaging. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, vol. 1, p. (2001). IEEE

Rahmann, S.: Polarization images: a geometric interpretation for shape analysis. In: Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, vol. 3, pp. 538–542 (2000). IEEE

Chen, G., He, L., Guan, Y., Zhang, H.: Perspective phase angle model for polarimetric 3d reconstruction. In: Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part II, pp. 398–414 (2022). Springer

Lu, J., Ji, Y., Yu, J., Ye, J.: Mirror surface reconstruction using polarization field. In: 2019 IEEE International Conference on Computational Photography (ICCP), pp. 1–9 (2019). IEEE

Goldstein, D.H.: Polarized Light, (2017)

Korger, J., Kolb, T., Banzer, P., Aiello, A., Wittmann, C., Marquardt, C., Leuchs, G.: The polarization properties of a tilted polarizer. Optics express **21**(22), 27032–27042 (2013)

Collett, E.: Field Guide to Polarization. Field Guide Series, (2005)

Bass, M., DeCusatis, C., Enoch, J.M., Lakshminarayanan, V., Li, G., MacDonald, C., Mahajan, V.N., Van Stryland, E.: Handbook of Optics, Third Edition Volume I: Geometrical and Physical Optics, Polarized Light, Components and Instruments(set). Handbook of Optics, (2009)

Cronin, T.W., Marshall, J.: Patterns and properties of polarized light in air and water. Philosophical Transactions of the Royal Society B: Biological Sciences **366**(1565), 619–626 (2011)

Nayar, S.K., Fang, X.-S., Boult, T.: Separation of reflection components using color and polarization. International Journal of Computer Vision **21**(3), 163–186 (1997)

Huynh, C.P., Robles-Kelly, A., Hancock, E.: Shape and refractive index recovery from single-view polarisation images. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1229–1236 (2010). IEEE

Jakob, W., Speierer, S., Roussel, N., Nimier-David, M., Vicini, D., Zeltner, T., Nicolet, B., Crespo, M., Leroy, V., Zhang, Z.: Mitsuba 3 Renderer. https://mitsuba-renderer.org

Wilkie, A.: The Advanced Rendering Toolkit. http://cgg.mff.cuni.cz/ART

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

Fatima, T., Pistellato, M., Torsello, A., Bergamasco, F.: One-shot hdr imaging via stereo pfa cameras. In: International Conference on Image Analysis and Processing, pp. 467–478 (2022).

Springer

Pistellato, M., Fatima, T., Wimmer, M.: Exploiting light polarization for deep hdr imaging from a single exposure. Sensors **23**(12), 5370 (2023)

Zhang, Z.: A flexible new technique for camera calibration. IEEE Transactions on pattern analysis and machine intelligence **22**(11), 1330–1334 (2000)

Strutt, J.W.: Xv. on the light from the sky, its polarization and colour. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **41**(271), 107–120 (1871)

Ichikawa, T., Purri, M., Kawahara, R., Nobuhara, S., Dana, K., Nishino, K.: Shape from sky: Polarimetric normal recovery under the sky. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 14832–14841 (2021)

Kupinski, M.K., Bradley, C.L., Diner, D.J., Xu, F., Chipman, R.A.: Angle of linear polarization images of outdoor scenes. Optical Engineering **58**(8), 082419 (2019)

Umow, v.N.: Chromatische depolarisation durch lichtzerstreuung. Phys. Z **6**, 674–676 (1905)

Mahmoud, A.H., El-Melegy, M.T., Farag, A.A.: Direct method for shape recovery from polarization and shading. In: 2012 19th IEEE International Conference on Image Processing, pp. 1769–1772 (2012). IEEE