



**HAL**  
open science

## Enhancing polarization 3D facial imaging: overcoming azimuth ambiguity without extra depth devices

Yudong Cai, Xuan Li, Fei Liu, Jiawei Liu, Kejian Liu, Zhiqiang Liu, Xiaopeng Shao

► **To cite this version:**

Yudong Cai, Xuan Li, Fei Liu, Jiawei Liu, Kejian Liu, et al.. Enhancing polarization 3D facial imaging: overcoming azimuth ambiguity without extra depth devices. *Optics Express*, 2023, 31 (26), pp.43891. 10.1364/OE.505074 . hal-04505664

**HAL Id: hal-04505664**

**<https://hal.science/hal-04505664>**

Submitted on 3 Apr 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.



L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



# Enhancing polarization 3D facial imaging: overcoming azimuth ambiguity without extra depth devices

YUDONG CAI,<sup>1,2</sup>  XUAN LI,<sup>1,2,\*</sup>  FEI LIU,<sup>1,2</sup>  JIAWEI LIU,<sup>3</sup>  
KEJIAN LIU,<sup>4</sup> ZHIQIANG LIU,<sup>1,2</sup> AND XIAOPENG SHAO<sup>1,2</sup>

<sup>1</sup>*School of Optoelectronic Engineering, Xidian University, Xi'an 710071, China*

<sup>2</sup>*Hangzhou Institute of Technology, Xidian University, Hangzhou 311200, China*

<sup>3</sup>*LAM, Aix Marseille Univ, CNRS, CNES, Marseille 13388, France*

<sup>4</sup>*Shanghai Aerospace Electronic Technology Institute, Shanghai 201109, China*

\*[lixuan@xidian.edu.cn](mailto:lixuan@xidian.edu.cn)

**Abstract:** Polarization 3D imaging has been a research hotspot in the field of 3D facial reconstruction because of its biosafety, high efficiency, and simplicity. However, the application of this technology is limited by the multi-valued problem of the azimuth angle of the normal vector. Currently, the most common method to overcome this limitation is to introduce additional depth techniques at the cost of reducing its applicability. This study presents a passive 3D polarization facial imaging method that does not require additional depth-capturing devices. It addresses the issue of azimuth ambiguity based on prior information about the target image's features. Specifically, by statistically analyzing the probability distribution of real azimuth angles, it is found that their quadrant distribution is closely related to the positions of facial feature points. Therefore, through facial feature detection, the polarized normal azimuth angle of each pixel can be accurately assigned to the corresponding quadrant, thus determining a precise unique normal vector and achieving accurate 3D facial reconstruction. Finally, our azimuth angle correction method was validated by simulated polarization imaging results, and it achieved accurate correction for over 75% of the global pixels without using additional depth techniques. Experimental results further indicate that this method can achieve polarization 3D facial imaging under natural conditions without extra depth devices, and the 3D results preserve edge details and texture information.

© 2023 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

## 1. Introduction

Three-dimensional (3D) reconstruction of faces plays a crucial role in various fields, such as security authentication [1], face media manipulation [2,3], face animation [4,5], and virtual reality. Generally, 3D face reconstruction methods can be classified into three categories: software-based, hardware-based, and image-based methods [6,7]. Among these, hardware-based and image-based methods are the most prevalent. For instance, LiDAR utilizes laser detection to derive the 3D shape, thereby achieving high precision and long detection distance [8]. However, it incurs high costs and poses potential risks to the eyes. Similarly, 3D imaging using structured light [9] enables accurate shape recovery but has challenges in dealing with long-range and complex reflectivity targets due to its reliance on striped light patterns. Compared to the above-mentioned active 3D reconstruction methods, passive shape recovery methods have great potential in various applications because of their biosafety. Binocular stereo vision [10], for example, mimics the way our eyes perceive 3D information using two cameras. However, the accuracy of reconstruction highly relies on the baseline, leading to reduced accuracy as the detection distance increases. Meanwhile, Similar to prevalent facial neural network technologies [11–13], image-based methods leverage one or multiple facial images for in-depth and comprehensive

abstract analysis of image data, offering advantages of simplicity and more realistic visual effects. Nonetheless, constrained by the input data, their applicability tends to be limited to specific demographics, often struggling to achieve high precision, while incurring high training and computational costs.

Face polarization 3D imaging has emerged as a prominent research area in the field of computation imaging and computer vision. By exploiting the polarization properties of light, it can acquire rich facial shape and texture information [14–16], enabling high-precision face reconstruction and recognition. The earliest research work was conducted by Wolff in the nineties, which focused on the capture, decomposition, and visualization of polarization images [17,18], even if there were older research works on shape recovery by polarization information since 1962 [19]. He established a model that used the reflected light of the object's surface to calculate 3D information for the first time [20]. This provided an essential theoretical foundation for polarization 3D imaging technology. O. Drbohlav et al. leveraged the polarization characteristic of diffuse reflected light to reconstruct objects' 3D shapes together with the SFS (Shape from Shading) method [21]. G. A. Atkinson et al. pointed out that using the polarization characteristic of diffuse reflected light can only determine the zenith [22]. Furthermore, K. P. Gurton accomplished the polarization 3D facial reconstruction in the long-wave infrared band by using the polarization characteristic of diffuse reflected light [23]. Later, A. Kadambi et al. confirmed the feasibility of the passive polarization 3D imaging technology based on experimental data. The polarization 3D imaging system of natural illumination was verified based on the "roughness and depth map" obtained by ToF [24]. Although much progress has been made in polarization 3D imaging, the imaging method with auxiliary equipment failed to make substantial breakthroughs in passive accurate 3D face reconstruction. There are major limitations, especially for complex ambient light under natural illumination.

This study proposes a passive 3D polarization facial imaging method without using depth-capturing devices. Firstly, the specific distribution pattern of azimuth angles for facial feature pixels in a quadrant is evaluated by probabilistic statistics. Then, by capturing polarization images, the azimuth angles and zenith angles of the pixel-level micro-surface normals are solved based on the polarization characteristic of diffusely reflected light from the target surface. To address the issue of azimuth ambiguity, this study exploits target prior feature information to construct azimuth templates. This provides conditional constraints to determine precise unique normals and avoid further distortions. Subsequently, accurate 3D facial reconstruction can be achieved. Finally, the imaging results indicate that, after applying the proposed method for correction, 75% of global pixels were accurately corrected. This is a breakthrough in passive technology that can produce a facial 3D shape while preserving edge detail and texture information.

## 2. Method

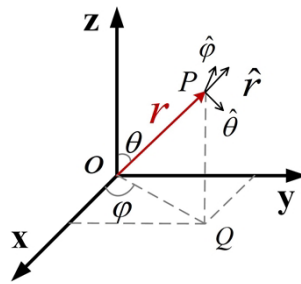
### 2.1. Acquiring accurate zenith and azimuth angles of human faces

To recover the 3D shape of a human face from a two-dimensional (2D) image, it is necessary to know the functional relationship between the 3D shape of each point on the face and the corresponding pixel in the 2D image. The proposed method operates in the coordinate system of the polarization camera and parameterizes the surface through an unknown depth function  $Z(x, y)$ , where  $(x, y)$  denotes the position in the polarization image. The 3D coordinates at  $(x, y)$  are determined by Eq. (1), where  $f$  denotes the focal length of the polarization camera in the  $x$  and  $y$  directions, and  $(x_0, y_0)$  denotes the principal point. For each captured polarized sub-image,

only the scaled Z-value of the true height for each pixel can be obtained.

$$P(x, y, z) = \frac{Z(x, y)}{f} \begin{bmatrix} x - x_0 \\ y - y_0 \\ f \end{bmatrix}. \quad (1)$$

According to the principle of spatial geometry and the definition of spherical coordinates shown in Fig. 1, any Cartesian coordinate point  $P(x, y, z)$  on a 3D surface in space can be represented with spherical coordinates  $P(r, \theta, \varphi)$ . In spherical coordinates,  $r$  denotes the distance between the origin  $O$  and  $P$ ,  $\theta$  denotes the angle between the directed line segment  $OP$  and the positive  $z$ -axis, and  $\varphi$  denotes the angle between the positive  $x$ -axis and the projection  $OQ$  of point  $P$  onto the  $x$ - $y$  plane observed in the counterclockwise direction.



**Fig. 1.** The sphere coordinates of a point  $P$  in space.

In the field of physics,  $\theta$  and  $\varphi$  in the spherical coordinate system are referred to as the zenith angle and azimuth angle, respectively. The corresponding relationships between  $x, y, z$  and  $r, \theta, \varphi$  are as follows:

$$\begin{aligned} x &= r \sin \theta \cos \varphi \\ y &= r \sin \theta \sin \varphi \\ z &= r \cos \theta. \end{aligned} \quad (2)$$

The human face is a completely closed surface and it can also be represented using spherical coordinates. Based on the derivation from calculus principles, for simplicity, this study assumes that the facial surface can be divided into infinitely continuous small-surface elements (which can be treated as points). Therefore, the convexity or concavity directions of each microsurface can be represented by its normal vector in spherical coordinates. In this study, under the assumption of orthographic projection, the visible part of the human face is represented by a depth function  $P = (x, y, Z(x, y))$ . The direction of the outward-pointing surface normal is defined as the cross product of the partial derivatives with respect to  $x$  and  $y$ . Then, the normal vector at point  $(x, y)$  can be represented as follows:

$$\vec{n} = \begin{bmatrix} -Z_x \\ -Z_y \\ 1 \end{bmatrix} = \begin{bmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{bmatrix} = \begin{bmatrix} \tan \theta \cos \varphi \\ \tan \theta \sin \varphi \\ 1 \end{bmatrix}. \quad (3)$$

where  $Z_x$  and  $Z_y$  denote the partial derivative of  $Z$  with respect to  $x$  and  $y$ , respectively.

The polarization-based 3D facial imaging model is constructed based on the spherical coordinate representation of the spatial positions described above. It should be noted that the

magnitude of  $\vec{n}$  is arbitrary, and only its direction matters, which allows eliminating common elements. Once the actual Z-values of the face are obtained from the laser scanner, the gradient field can be derived using differentials. Conversely, with knowledge of the gradient field, the height can also be recovered through integration. The relationship between the height, the gradient field ( $p, q$ ), and the zenith and azimuth angles, is shown below:

$$Z_{real}(x, y) \begin{matrix} \xrightarrow{\text{derivation}} \\ \xleftarrow{\text{integration}} \end{matrix} \left\{ \begin{array}{l} p = \partial Z_{real} / \partial x = \tan \theta \cos \varphi \\ q = \partial Z_{real} / \partial y = \tan \theta \sin \varphi \end{array} \right. \quad (4)$$

According to the physical model, since only half of the frontal face in the photograph can be captured, the zenith angle is normally constrained to be  $[0, 90)$ , as shown in Fig. 2, and  $\varphi \in [0, 360]$ . Therefore, the value range of  $\varphi_{Cos}$  is determined only by the sign of  $p$ , which restricts  $\varphi_{Cos}$  to the left or right side of the y-axis, as shown in the range illustrated in Fig. 3(a). Similarly, the range of  $\varphi_{Sin}$  is determined only by the sign of  $q$ , which limits  $\varphi_{Sin}$  to the upper or lower side of the x-axis, as shown in the range in Fig. 3(b).

$$\theta = \arctan \left( \sqrt{p^2 + q^2} \right) = \arctan \left( \sqrt{(\partial Z_{real} / \partial x)^2 + (\partial Z_{real} / \partial y)^2} \right) \quad (5a)$$

$$\left\{ \begin{array}{l} \varphi_{Cos} = \arccos(p / \tan \theta) = \arccos(\partial Z_{real} / \partial x / \tan \theta) \\ \varphi_{Sin} = \arcsin(q / \tan \theta) = \arccos(\partial Z_{real} / \partial y / \tan \theta) \end{array} \right. \quad (5b)$$

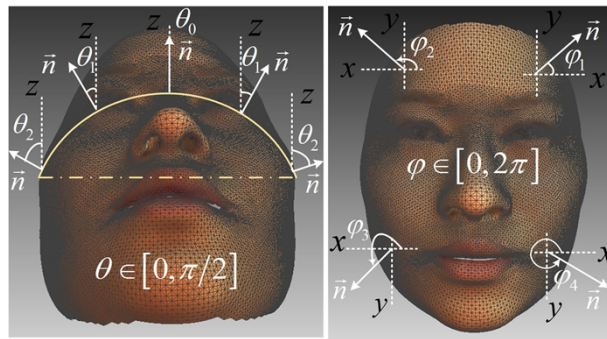


Fig. 2. Illustration of normal zenith and azimuth angle ranges.

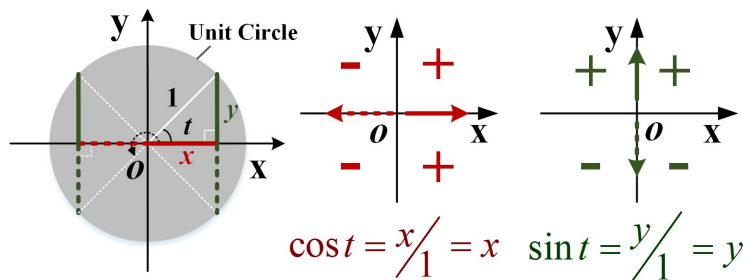


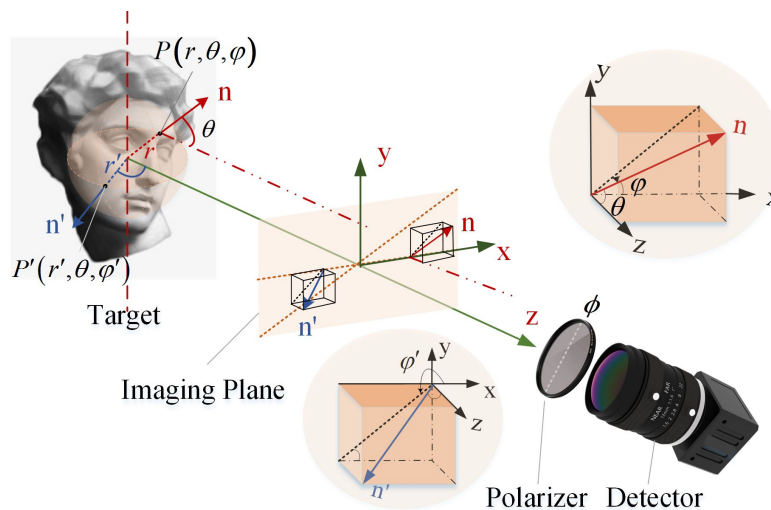
Fig. 3. The signs of Sine and cosine functions in different quadrants.

Therefore,  $\varphi_{Cos}$  and  $\varphi_{Sin}$  can be obtained based on Eq. (5(b)), which provides two facial azimuth angle images. Then, we can search for points in the  $\varphi_{Sin}$  image where  $\varphi_{Cos}$  is less than or greater than 0. For the point where  $\varphi_{Cos}$  is less than 0,  $2\pi$  is used to subtract  $\varphi_{Sin}$  to obtain the true azimuth angle for this point. For the point where  $\varphi_{Cos}$  is greater than or equal to 0,  $\varphi_{Sin}$  is directly used as the true azimuth angle for this point. In this approach, the true azimuth angle of each pixel in the facial polarization image can be obtained. The results demonstrate that the azimuthal distribution of the entire region has a certain regularity and can be determined through feature point segmentation.

$$\begin{cases} \varphi = \varphi_{Cos}, & \text{if } \varphi_{Cos} * \varphi_{Sin} > 0 \\ \varphi = 2\pi - \varphi_{Sin}, & \text{if } \varphi_{Cos} * \varphi_{Sin} < 0 \end{cases} \quad (6)$$

## 2.2. Polarization 3D imaging model for human faces

According to the Fresnel formula and the polarization properties of light waves, the polarization information can be extracted from the light-intensity image on the imaging plane of the detector, and this information is closely related to the surface shape of the target object. The polarization-based 3D facial imaging model inverts the normal information of each point on the face surface by exploiting the spatial geometric relationship between the polarization information in the light waves emitted from the facial surface and the normal of its surface. The imaging model is illustrated in Fig. 4.



**Fig. 4.** Polarization 3D face imaging model schematic.

Assuming that the facial surface is an ideal Lambertian surface, the light rays emitted from this surface undergo diffuse reflection and multiple refractions inside the face. In Fig. 4, light is reflected from the surface and propagates along the positive z-axis, and two points  $P$  and  $P'$  on the surface correspond to the normals  $n$  and  $n'$ , respectively. The zenith angle  $\theta$  denotes the refracted angle of the outgoing light wave,  $x$  and  $y$  denote the 2D coordinates on the imaging plane, and  $\varphi$  denotes the angle between the projection of the facial normal on the imaging plane and the positive x-axis, also known as the normal azimuth angle. Notably, two different normals can pass through the same polarizer at the same polarization angle. This leads to an inherent issue in polarization 3D imaging known as the normal azimuth angle ambiguity.



In fact, obtaining the features of any real human face requires an accurate calculation of the normal of each facial pixel. As analyzed in Section 2.1, solving the normal of each point on the facial surface is equivalent to solving the zenith angle  $\theta$  and azimuth angle  $\varphi$  corresponding to each normal. The polarization information of the outgoing light wave can be calculated by using the intensity facial image obtained by the detector, and based on this polarization information, the calculation of  $\theta$  and  $\varphi$  can be achieved.

In the setup shown in Fig. 4, a linear polarizer is placed in front of the detector, and it can be rotated at any angle to obtain the polarization information of the outgoing light wave. After choosing any direction as the initial direction of the linear polarizer, the polarizer is rotated, and the variation of light wave intensity information with the rotation angle is represented as:

$$i_{\theta_j}^{\text{mod}}(i_{\text{un}}, \phi, \rho) = i_{\text{un}}(1 + \rho \cos[2\theta_j - 2\phi]). \quad (7)$$

The sinusoid has a period of  $\pi$  and is characterized by three quantities that are together referred to as a polarization image [25,26]. The unpolarized intensity,  $i_{\text{un}}$ , is the mean value of the sinusoid. This is the intensity that would have been observed without a polarizer, and thus it depends on the reflectance properties of the surface and the illumination in the scene. The phase angle,  $\phi \in [0, \pi]$ , defines the phase shift, which is directly related to the angle of the linearly polarized component of the reflected light and can be defined as the angle of maximum or minimum transmission. The degree of polarization (DoP),  $\rho \in [0, 1]$ , is the ratio between the amplitude and mean value of the sinusoid. The three components of a polarization image depend on the local surface geometry at the point of reflection and the material properties. A simplified expression can be obtained by taking ratios between different polarizer orientations:

$$\frac{i_{\theta_j}^{\text{mod}}(i_{\text{un}}, \phi, \rho)}{i_{\theta_k}^{\text{mod}}(i_{\text{un}}, \phi, \rho)} = f_{\theta_j, \theta_k}(\phi, \rho) = \frac{1 + \rho \cos[2\theta_j - 2\phi]}{1 + \rho \cos[2\theta_k - 2\phi]}. \quad (8)$$

This can remove any dependency on  $i_{\text{un}}$  and thus on any assumed reflectance model, material properties, or illumination. Therefore, using this ratio expression alone does not require estimating albedo and lighting, nor does it require assuming an underlying reflectance model [25].

In the actual solution process, the Stokes parameters  $[S_0, S_1, S_2, S_3]^T$  can be obtained by rotating the polarizer at multiple polarization angles to obtain a set of images [27,28]. Specifically,  $S_0$  denotes the power of the incident beam,  $S_1$  and  $S_2$  respectively denote the power of 0 and 45 linear polarization, and  $S_3$  denotes the power of right circular polarization. For unpolarized natural light, the value of  $S_3$  is small and is often ignored. With the calculated Stokes parameters, the values of DoP and AoP (the angle of polarization) can be obtained.

$$\begin{cases} \rho = \sqrt{S_1^2 + S_2^2}/S_0 \\ \phi = \frac{1}{2} \arctan_2(S_2, S_1) \end{cases}. \quad (9)$$

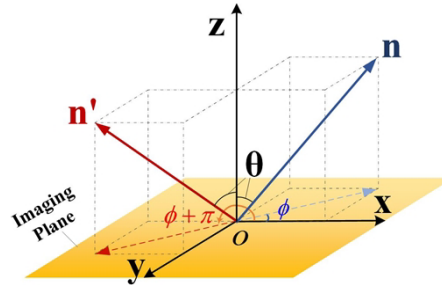
where  $\arctan_2$  denotes the four-quadrant arctangent operator [29].

Polarization images can be used to constrain the orientation of the surface normal at each pixel. The exact nature of the constraint relies on the polarization model used. Using the diffuse polarization model, the phase angle is the polarizer angle at which the maximum brightness is observed. It determines the azimuth angle  $\varphi$  of the surface normal in the range  $[0, 2\pi]$ , with an ambiguity of  $\pi$ :  $\varphi = \phi$  or  $\varphi = \phi + \pi$ . Meanwhile, the DoP  $\rho$  is related to the refractive index  $n$  and the zenith angle  $\theta$  of the surface normal in the observer-centered coordinates (i.e., the angle between the normal and the observer, which falls in the range  $[0, \pi/2]$ ), and the relationship is

represented below [30]:

$$\cos \theta = \vec{n} \cdot \vec{v} = f(\rho, n) = \sqrt{\frac{2\rho + 2n^2\rho - 2n^2 + n^4 + \rho^2 + 4n^2\rho^2 - n^4\rho^2 - 4n^3\rho\sqrt{-(\rho-1)(\rho+1)} + 1}{n^4\rho^2 + 2n^4\rho + n^4 + 6n^2\rho^2 + 4n^2\rho - 2n^2 + \rho^2 + 2\rho + 1}}. \quad (10)$$

As shown in Fig. 5, two normal vectors with different orientations have the same AoP value. When the polarizer completes one full rotation ( $2\pi$ ), two pieces of maximum brightness are observed, resulting in  $\pi$  ambiguity. To overcome the uncertainty introduced by the azimuthal ambiguity of the surface normal, it is necessary to apply constraints to satisfy specific restrictions.



**Fig. 5.** The schematic of normal azimuth angle ambiguity

Once accurate zenith and azimuth angles are obtained from polarization sub-images, the surface normal of the target can be represented. To address the issue that the normal vectors derived from polarization features cannot directly depict the 3D shape of the target, Frankot and Chellappa developed an enhancement of the integrability method [31] to establish the correspondence between normal vector parameters and the 3D shape of the object. The expression for the surface function  $Z$  in terms of the gradient field  $(p, q)$  is given below:

$$Z = F^{-1} \left\{ \frac{j}{2\pi} \frac{u/M F\{p\} + v/N F\{q\}}{(u/M)^2 + (v/N)^2} \right\}. \quad (11)$$

where  $F$  and  $F^{-1}$  represent the discrete Fourier transform and inverse discrete Fourier transform, respectively.  $(M, N)$  denotes the height and width of the target image, respectively.  $(u, v)$  represents frequency domain coordinates, with values ranging from  $(-M/2, -N/2)$  to  $(M/2, N/2)$ .

### 2.3. Azimuth angle correction and facial feature probability patterns

In Section 1, the analysis demonstrates that for front-facing real human faces, there is a well-defined range for the distribution of normal azimuth angles. Therefore, if image processing or neural network-based face detection techniques can be used, and with the assistance of existing mature feature point detection technology, the location of specific parts of the face in an image can be accurately obtained. To investigate the relationship between feature points and azimuthal distribution, the feature points around each pixel are grouped to calculate the probability of the pixel belonging to each region. Suppose that there are  $k$  regions, where the  $i$ -th region has  $m_i$  feature points. For each pixel  $t$ , suppose that there are  $n$  feature points around it as  $s_1, s_2, \dots, s_n$ .



Then, the probability that  $t$  belongs to the  $i$ -th region  $P_i(t)$  can be calculated as follows:

$$P_i(t) = \frac{1}{n} \sum_{j=1}^{m_i} \omega(\|t - s_i\|). \tag{12}$$

where,  $\|\cdot\|$  denotes the Euclidean distance;  $\omega(\|t - s_i\|)$  denotes the weight at the distance  $\omega(\|p - q_j\|) \propto \|t - s_i\|$ , and it can generally be calculated using a Gaussian function:

$$\omega(\|t - s_i\|) = e^{-\frac{\|t - s_i\|^2}{2\sigma^2}}. \tag{13}$$

where  $\sigma$  denotes the standard deviation of the Gaussian function, and  $\sigma$  is set to 1.6 in this study.

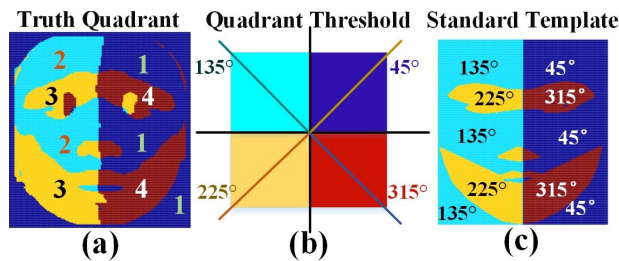
By calculating the probability of each pixel belonging to each region in the face image, pixels can be assigned to their corresponding regions, and the relationship between feature points and their corresponding regions can be determined. Thus, when processing new data or new images, the pixels of the feature points can be simply attributed to the corresponding quadrant according to the facial feature recognition results. Combined with the discussion in Section 2.1, any captured frontal face image can be used for feature point detection to obtain the range of azimuth distribution at each pixel. In this study,  $I_0$  is taken as the processing target. After pre-processing (e.g., denoising), the pure YOLO series face detection algorithm is employed to retrieve facial feature coordinates. Then, the facial feature points are clustered to obtain several clusters of feature points.

The clustering approach is described as follows:

Let the set of facial feature points be  $P$ , and it needs to be divided into four clusters. Choose  $k$  points as initial centroids (randomly or specific feature points).

- 1) For each feature point, calculate its distance to the  $k$  centroids and assign it to the cluster with the closest distance.
- 2) For each cluster, recalculate its centroid. Repeat steps 1) and 2) until the clusters no longer change or the predetermined number of iterations is reached.

Through clustering, the face is divided into distinct regions using pre-defined prior information. Specifically, for a polarized facial image, the background region to the right of the vertical midline is denoted as the 1st quadrant, while the background region to the left of the vertical midline is referred to as the 2nd quadrant. The eye and cheek regions to the left of the vertical midline are referred to as the 3rd quadrant, while the eye and cheek regions to the right of the vertical midline are referred to as the 4th quadrant, as shown in Fig. 6(a).



**Fig. 6.** Visual depiction of the facial azimuth quadrant distribution: (a) Quadrant distribution after clustering truth azimuths; (b) Threshold illustration for each quadrant; (c) Generated standard template for azimuth correction.

To facilitate the subsequent calculation process, azimuth angle thresholds are set for each region according to the azimuth angle quadrant distribution, as illustrated in Eq. (14). Thus, the

azimuth angle thresholds for the first, second, third, and fourth quadrants are set to 45°, 135°, 225°, and 315°, respectively, as shown in Fig. 6(b). In addition, in Fig. 6(a), the azimuth angles for each pixel in Regions 1, 2, 3, and 4 lie within the ranges [0, 90), [90, 180), [180, 270), and [270, 360), respectively. By setting the azimuth angle thresholds in this way, the standard azimuth angle map can be obtained, as shown in Fig. 6(c). The reason behind setting this threshold is to facilitate subsequent identification of failed pixels in the statistical constraints. Pixels deviating from my template by over 45 degrees are considered to be outside the same quadrant, indicating constraint failure.

$$\varphi_{st} = \frac{\pi}{2} \text{Num\_Reg} - \frac{\pi}{4} \quad (14)$$

where  $\text{Num\_Reg}$  refers to the number of quadrant to which the pixels in the face area belong.

After obtaining the standard azimuth angle threshold map corresponding to the target image, the difference between the computed polarization azimuth angle and the standard azimuth angle is calculated. If the absolute value of the angle difference exceeds the preset threshold ( $\pi/4$ ), the polarization azimuth angle for the point is corrected by 180°; otherwise, the polarization azimuth angle for the point is directly considered as the accurate azimuth angle. In other words, for the azimuth angle calculated using polarization information, if its value is closer to the azimuth angle threshold for the point, it indicates that the azimuth angle estimated by Eq. (9) is accurate, and thus accurate micro-surface normal information is obtained by directly using the polarization characteristics of the reflected light; Otherwise, the azimuth obtained by solving using the polarization characteristics will be corrected to achieve a 180° flip.

$$\Lambda_{sgn} = \arg \min_{\Lambda_{sgn}} \|\psi_{pl} + \Lambda_{sgn}\pi - \psi_{std}\|_2^2, \quad \Lambda_{sgn} \in [0, 1]. \quad (15)$$

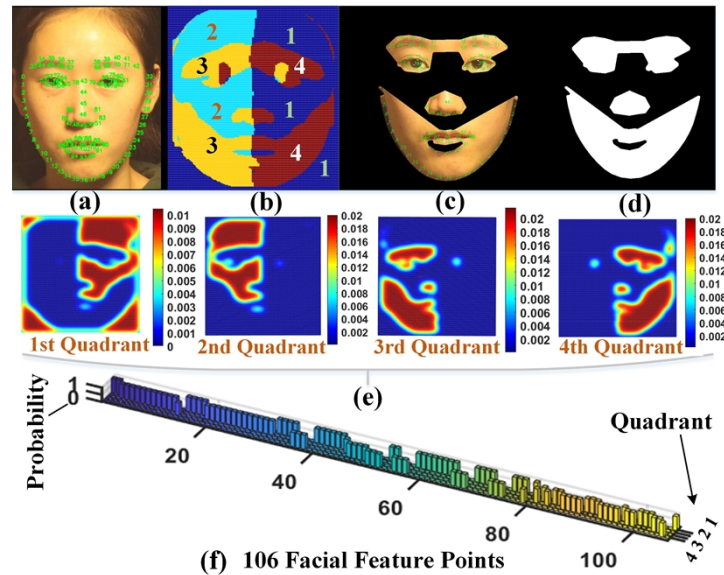
where  $\psi_{pl}$  denotes the polarization azimuth angle of each pixel calculated using Eq. (9), and  $\psi_{std}$  denotes the standard azimuth angle threshold corresponding to each pixel. Finally, with the corrected azimuth angles and zenith angles, the normal vectors at each pixel on the face surface can be calculated using the proposed polarization 3D face imaging model. These normals can be used for precise 3D facial reconstruction.

### 3. Experimental results and discussion

#### 3.1. Facial feature detection and location probability analysis

In this section, experiments are designed and conducted to validate the effectiveness and accuracy of each theory proposed. First, the operational capability of the proposed method for the azimuth angle correction is verified. By using real 3D facial data, an accurate and precise azimuth angle distribution is obtained, as illustrated in Fig. 7(b). Meanwhile, the YOLO series facial detection algorithm is adopted to retrieve 106 key points for the same facial target, as demonstrated in Fig. 7(a). Subsequently, following the method proposed in Section 2.3, the probability that each feature point falls within a specific quadrant is calculated, as illustrated in Figs. 7(e1)-7(e4). These figures represent the probability distribution of each pixel within the facial region falling within one of the four quadrants.

Figure 7(f) displays the calculated probability of the 106 feature points falling within each quadrant specifically. Through experimental observation, it can be found that feature points 1 to 3, 30 to 33 had a higher probability of falling within the first quadrant; features points 44 to 47 had a higher probability of falling within the second quadrant; features points 4 to 17, 34 to 38, 48 to 58, 65 to 68, 73 to 75, 79, 81, 83, 85 to 88, 95 to 99, and 104 to 105 had a higher probability of dropping in the third quadrant; features points 18 to 29, 39 to 43, 59 to 64, 69 to 72, 76 to 78, 80, 82, 84, 89 to 94, 100 to 103, and 106 are more probable to drop in the fourth quadrant. Therefore, the region can be concatenated and divided into segmented regions as shown in Fig. 7(c). So

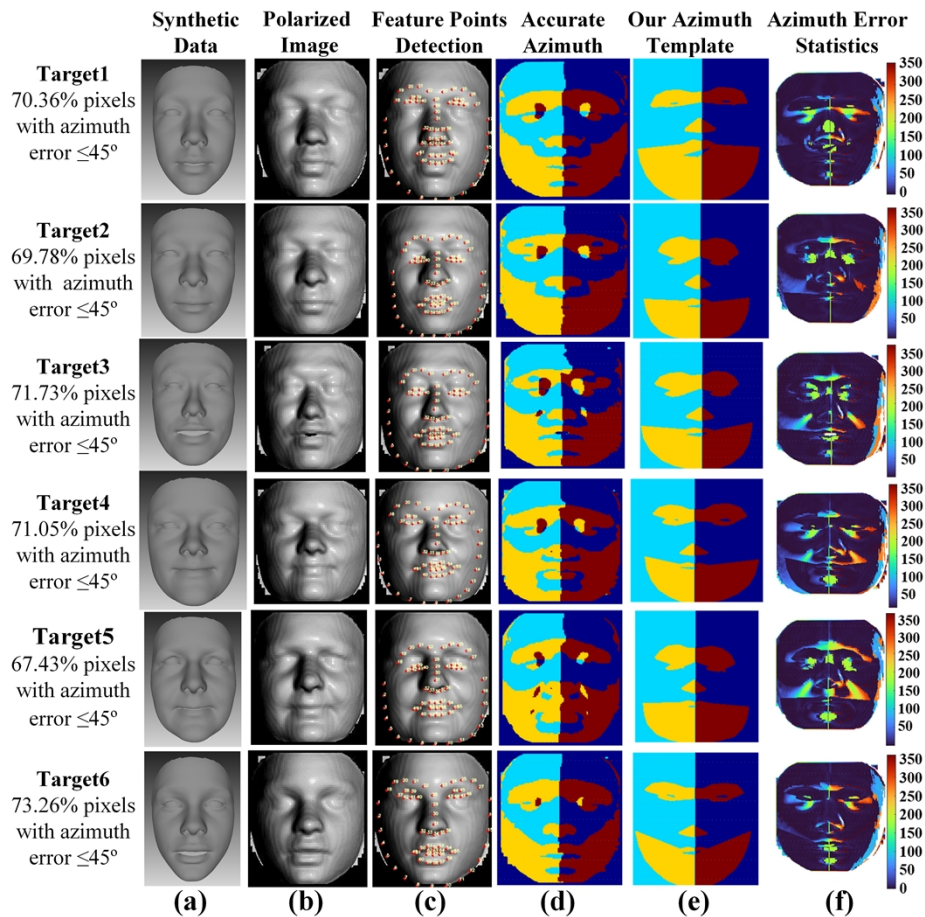


**Fig. 7.** Facial feature detection and the probability of each feature point falling within a certain quadrant: (a) Frontal face image and the location of its feature points; (b) The quadrant distribution of the accurate azimuth angle; (c) Segmentation of the face region using the detected feature points; (d) Schematic of the segmented face region; (e) Pixel probability within specific quadrants in the image; (f) Probability of 106 facial landmarks falling within each quadrant.

far, for any given image, the azimuth region segmentation corresponding to the image can be obtained through facial feature point detection.

To further validate the effectiveness of polarization angle correction using segmented azimuth templates, six 3D facial models were generated by software, as shown in Fig. 8(a). According to the theory presented in Section 2.1, accurate azimuth angles and their quadrant distributions were computed for each model, as illustrated in Fig. 8(d). Following the Blinn-Phong reflection model, the generated 3D data was further used to simulate the polarization sub-images, as illustrated in Fig. 8(b). The feature point distributions detected from the 0-degree polarized image using the pure YOLO algorithm are shown in Fig. 8(c). By applying the proposed clustering method based on the feature point positions, the azimuth quadrant template was obtained and shown in Fig. 8(e). Subsequently, a statistical analysis was conducted on the azimuth angle errors between the corrected azimuth angles and the accurate azimuth angles, as illustrated in Fig. 8(f).

It can be seen from Fig. 8 that the accuracy of feature point identification has a great impact on the azimuth correction results. As indicated in Fig. 8(c), due to inaccurate feature point localization, the azimuth error distribution of Target 2 and Target 5 exhibits a richer color spectrum, where different colors represent varying degrees of azimuth angle error. Consequently, the proportion of pixels with corrected azimuth errors under 45-degrees accounts for only 69.78% and 67.43% of the total pixels for these two targets, indicating their respective correctly corrected pixels ratios. However, for the other four targets, namely, Target 1, Target 3, Target 4, and Target 6, the proportion of pixels with azimuth errors below 45 degrees surpasses 70%. Notably, Target 6 exhibits an even higher percentage of pixels (73.26%) with errors below 45 degrees. Across all targets, the similarity in the distribution of the two templates for Target 6 is prominent, as illustrated by a comparison between Fig. 8(d) and 8(e).

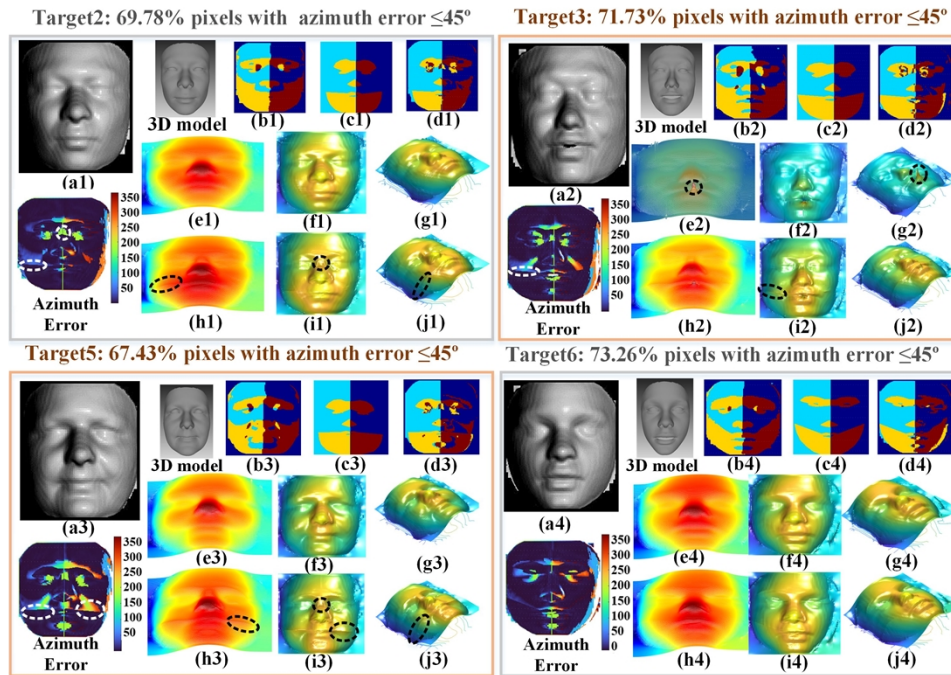


**Fig. 8.** Qualitative evaluation of face normal azimuth correction using target prior information: (a) Software-generated 3D human face model; (b) Simulated 0-degree polarized image using the generated 3D data based on the Blinn-Phong reflection model; (c) Feature point distribution obtained by applying the pure YOLO algorithm to 0-degree polarized images; (d) Accurate azimuthal quadrant template derived from depth; (e) Azimuthal quadrant template derived from the proposed method; (f) Statistical analysis of the azimuth error between the azimuth corrected by the proposed method and the accurate azimuth angle.

### 3.2. Analysis of simulated data for polarization 3D facial imaging results

Regarding the previously mentioned six sets of simulated data, a comparative analysis was conducted focusing on the targets with lower recognition rates, namely, Target 2, Target 3, Target 5, and Target 6. The accurate azimuth quadrant template, as illustrated in Fig. 9(b\*), was derived from the depth following Eqs. (5(b)) and (6). Fig. 9(d\*) presents the results marked with quadrant counts for all corrected azimuths, where different colors represent different quadrants. Figs. 9(e\*) and 9(h\*) display the depth distribution of the 3D results obtained using accurate azimuths versus the corrected azimuths using the proposed method, where different colors represent different heights. Figures 9(f\*) and 9(g\*) depict the 3D results reconstructed from the accurate azimuth and zenith angle in different viewpoints, and similarly, Figs. 9(i\*) and 9(j\*) show the 3D results reconstructed from the zenith and the corrected azimuth angles through the proposed method in two different perspectives.

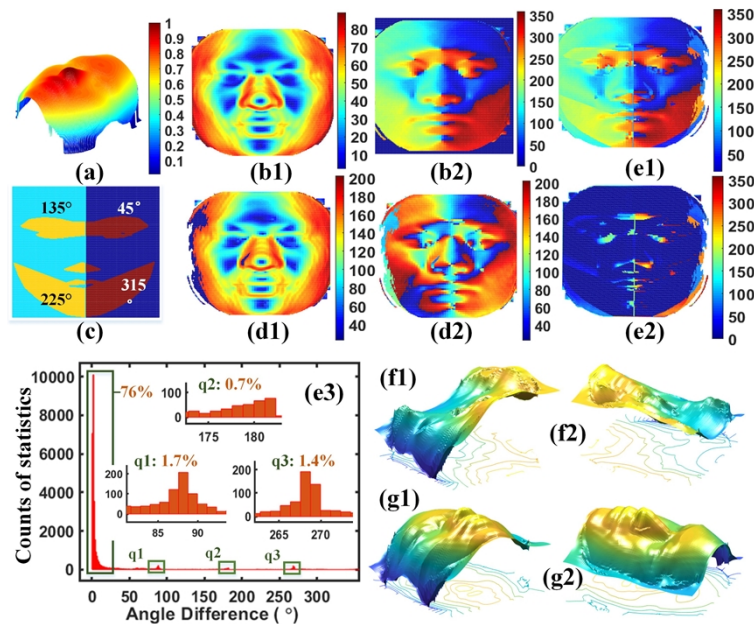




**Fig. 9.** Polarization 3D imaging of four sets of simulated facial data: (a\*) Simulated 0-degree polarized image using the generated 3D data based on the Blinn-Phong reflection model; (b\*) Accurate azimuthal quadrant template derived from depth; (c\*) Azimuthal quadrant template derived from proposed method; (d\*) Corrected azimuthal quadrant template; (e\*) and (h\*) show the depth maps of the 3D results obtained using the accurate azimuth versus the corrected azimuth by the proposed method; (f\*) and (g\*) illustrate the two-view 3D results reconstructed from accurate azimuth. (i\*) and (j\*) display the two-view 3D results reconstructed after azimuth correction using the proposed method.

Figure 9 illustrates that the accuracy of feature point identification not only affects the azimuth correction but also significantly impacts the polarized 3D reconstruction results. In the processing results for Target 5, the proportion of corrected azimuth error pixels under 45 degrees is only 67.43%, it's noticeable that distortion and concavity coexist in the reconstructed results shown in Figs. 9(h3), 9(i3), and 9(j3). When compared with the accurate azimuth quadrant template in Fig. 9(b3), it is evident that the size of the cheek region in Fig. 9(c3) is notably smaller than the accurate value, and there are variations in the eyebrow area. This is the same reason that in the reconstruction result of Target 2 in Fig. 9(i1), there also exists a concavity in the eyebrow bone region. However, in the case of Target 6 with more accurate segmentation results, the above pronounced distortions no longer exist. It is worth noting that there are obvious small bumps near the mouth area in Figs. 9(e2), 9(f2), and 9(g2). This is due to the simulated polarization sub-images from the software-generated 3D model, where the mouth region was assigned zero depth because the target is in the state of grinning and laughing. Consequently, due to multiplication and division processes, the calculated polarization information is infinitely amplified, resulting in the observation of small protrusions near the mouth area. It is reasonable to believe that with more accurate feature detection, the proposed method can further reduce azimuthal angle errors and improve the three-dimensional reconstruction accuracy to a greater extent.

Furthermore, to delve deeper into the performance of the proposed method in the polarization information interpretation process, an analysis of zenith and azimuth angles was conducted using

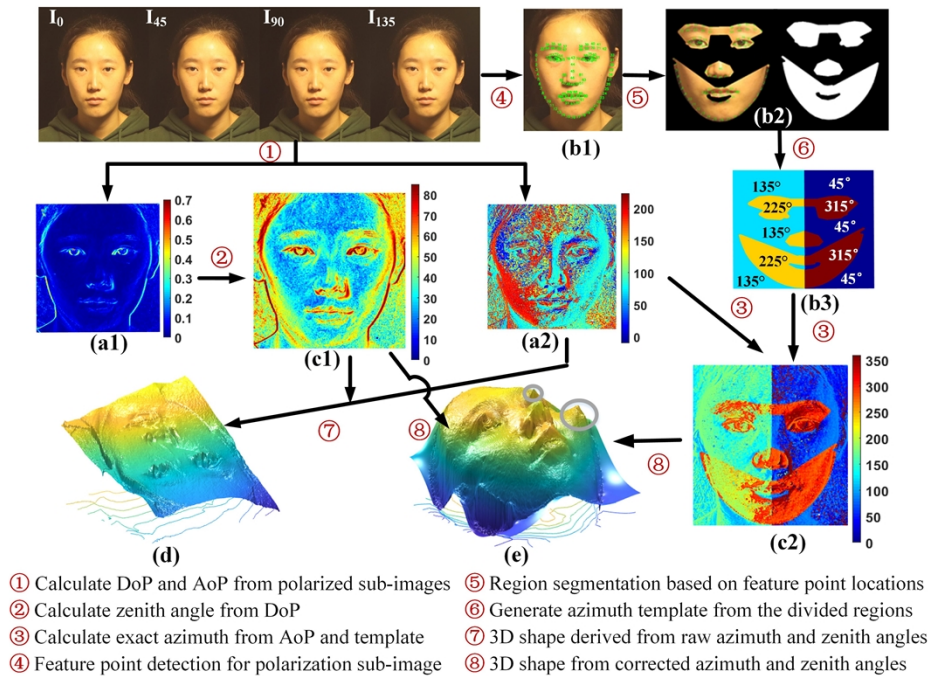


**Fig. 10.** Overall validation results of the proposed polarization 3D imaging method using simulated polarization data: (a) Depth distribution of the human face to be imaged; (b1) and (b2) show the accurate zenith and azimuth angles calculated from the depth, respectively; (c) The azimuth quadrant constraint template obtained using the proposed method; (d1) and (d2) show the zenith and ambiguous azimuth angles calculated from the simulated polarization images, respectively; (e1) The distribution of the corrected azimuths by the proposed method; (e2) and (e3) The distribution of the difference between the corrected azimuth distribution and the accurate azimuths, and the statistical histogram distribution of the difference, respectively; (f1) and (f2) show the 3D results reconstructed directly using the zenith angle and the ambiguous azimuthal angle in different viewpoints; (g1) and (g2) show the 3D results reconstructed using the zenith angle and the corrected azimuthal angle in different viewpoints, respectively

the simulated data from Fig. 8. The exact height distribution is presented in Fig. 10(a). Using the approach from Method 1, solving for the partial derivative of height yielded the normal distribution, and then accurate zenith and azimuth angles were obtained and shown in Figs. 10(b1) and 10(b2), respectively. Then, polarization sub-images corresponding to each polarizer were derived by using the Blinn-Phong reflection model. With these polarized sub-images available, the zenith angle and the AoP were calculated using the polarization model from Method 2, as shown in Figs. 10(d1) and 10(d2). A comparison with Figs. 10(b\*) reveals that the calculated zenith angle distribution from polarization information generally matches the actual situation, whereas the azimuth angle distribution exhibits poor accuracy.

To validate the accuracy of the constrained azimuth, the angle difference distribution shown in Fig. 10(e2) was calculated by making a difference with the accurate azimuth. Through analyzing the distribution information statistically, as depicted in Fig. 10(e3), it can be observed that the majority of azimuth angles have been corrected, and only a few exhibit significant errors. Specifically, the percentage of pixels with corrected azimuth errors under 45 degrees reaches 75%. Meanwhile, there are three peaks in the range of 80-95, 170-185, and 260-275 degrees, with a percentage of pixels of 1.7%, 0.7%, and 1.4%, respectively. The reason behind the majority of azimuth errors being within 50 degrees is that the azimuth correction criterion restricts the difference from the standard value should not exceed 45 degrees. This indicates that most azimuth





**Fig. 11.** The validation results of the proposed polarization 3D imaging method using actual captured polarization data: (a1) and (a2) show the polarization degree and polarization angle calculated from the acquired polarization sub-images; (b1)-(b3) show the results of obtaining the azimuthal quadrant constraint template by using the method proposed in this paper; (c1) and (c2) show the zenith angle and the corrected azimuth angle, respectively; (d) and (e) show the reconstructed results obtained by using ambiguous azimuths and the corrected azimuths from the proposed method, respectively.

angles have been adjusted to fall within the same quadrant as the standard value. Figs. 10(f1) and 10(f2) display the 3D results reconstructed directly based on zenith angle and uncorrected azimuth angle. Due to the ambiguity of the azimuth angle, the reconstructed results seem flat and distorted and lack any distinct features of the target. By using the azimuth template shown in Fig.12(c), the corrected 3D results are obtained and illustrated in Figs. 10(g1) and 10(g2). At this stage, the results exhibit facial morphology features and are relatively ideal.

### 3.3. Analysis of captured images for polarization 3D facial imaging results

After analyzing the processing results of a large volume of simulated polarization data, it becomes evident that the method proposed in this paper has high feasibility and robustness. To validate the performance of the proposed method in real-world scenarios, following the imaging model illustrated in Fig. 4, the target was imaged under natural lighting at a distance of about 5 m from the camera. Four polarization sub-images of the target face were captured using a Thorlabs LPVISC100-MP2 linear polarizer placed in front of a Canon EOS 77D camera (the resolution is  $6000 \times 4000$  pixels, the exposure time is  $1/40$  s, and the focal length is 35 mm). Subsequently, the proposed method was applied to process the captured images, thereby generating azimuth templates and reconstructing the results, as shown in Fig. 11.

Figures 11(a1) and 11(a2) respectively illustrate the calculated DoP and AoP values using Eq. (9) from the acquired polarization sub-images. Figures 11(b1)-11(b3) demonstrate the process of obtaining the azimuth quadrant constraint template from the 0-degree polarizer image using the

proposed method. This corresponds to steps 4, 5, and 6 in Fig. 11, specifically involving feature detection on the 0-degree polarization sub-image, segmentation based on the location of the feature points using the clustering method described in Section 2.3, and eventually generating a standard azimuth template based on the segmented regions. Figure 11(c1) presents the calculated zenith angle following Eq. (10), while Fig. 11(c2) depicts the azimuth angle corrected using the template shown in Fig. 11(b3), as proposed in this paper.

The processing of real-world data can be divided into two routes as shown in Fig. 11. The first route involves obtaining the azimuth standard template through feature point detection and facial region segmentation. The other route involves directly calculating the DoP and AoP values from the polarization sub-images and then decoding the zenith angles and azimuth angles. Subsequently, the standard azimuth template obtained from the first route is used to constrain the azimuth angles calculated through polarization sub-images. Then, the corrected azimuth angles and zenith angles are used to calculate the facial normal, and the final 3D results are obtained using the FC algorithm [31], as illustrated in Fig. 11(e). The 3D results obtained without azimuth angle correction are demonstrated in Fig. 11(d). Compared to the processing results for simulated data, this result exhibits distortion in the cheek region and slightly worse performance. However, the processing results from real captured polarization sub-images indicate that the proposed method can accomplish polarization 3D imaging of human faces under natural conditions, without requiring additional depth capture devices; meanwhile, it preserves the details and edge information in the 3D results. The slightly reduced performance is due to the limitations in accurately capturing polarization information using the current acquisition methods. Additionally, the operation steps in this process indicate that the passive and purely physical nature of this method makes it feasible for rapid, simple, and low-cost 3D facial imaging products.

#### 4. Conclusion

This paper proposes a passive 3D polarization imaging method for reconstructing 3D facial surfaces under natural conditions without requiring assistance from other depth techniques. Compared to traditional polarization-based 3D imaging methods, the proposed method adopts the prior feature information of facial polarization images to address the ambiguity of the azimuth angle. By simulating the polarization sub-images from real 3D facial models and obtaining the AoP and accurate azimuth angles, the probability distribution of the real azimuth angles is statistically analyzed, and it is found that their distribution quadrants are closely related to the locations of facial feature points. Therefore, through facial feature detection, the azimuth angles of each pixel's polarized normal vector can be correctly classified into the corresponding quadrants, thus obtaining a standardized azimuth angle distribution. Then, this distribution is exploited to constrain the azimuth angles calculated from AoP to determine the accurate unique normal vector, thereby achieving precise polarization-based 3D facial reconstruction. Finally, simulation results of polarization imaging indicate that the proposed method can correct polarized azimuth angles, achieving accurate correction for over 75% of the global pixels without extra depth devices. With more accurate feature detection, the corrected azimuth errors will further decrease, leading to improved 3D reconstruction precision. Additionally, more generalized correction methods can be utilized to optimize global corrections while maintaining local constraints, thereby improving overall results. All experiments demonstrate that the proposed approach outperforms existing methods in terms of visual quality, implementation cost, and computational efficiency. Furthermore, the 3D results effectively preserve edge details and texture information.

**Funding.** National Natural Science Foundation of China (62205256); Shanghai Aerospace Science Innovation Foundation (SAST2022-094); CAS Key Laboratory of Space Precision Measurement Technology (SPMT2023-02).

**Acknowledgment. Informed Consent Statement.** The human faces displayed in this manuscript are from virtual simulations and the fourth author Jiawei Liu. The author has confirmed that informed consent was obtained.

**Disclosures.** The authors declare no conflicts of interest.

**Data availability.** Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

## References

1. M. Uzair, A. Mahmood, F. Shafait, *et al.*, “Is spectral reflectance of the face a reliable biometric?” *Opt. Express* **23**(12), 15160–15173 (2015).
2. V. Blanz and T. Vetter, “A morphable model for the synthesis of 3D faces,” in *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, (1999), 187–194.
3. J. Thies, M. Zollhofer, M. Stamminger, *et al.*, “Face2face: Real-time face capture and reenactment of rgb videos,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, (2016), 2387–2395.
4. C. Cao, Y. Weng, S. Lin, *et al.*, “3D shape regression for real-time facial animation,” *ACM Trans. Graph.* **32**(4), 1–10 (2013).
5. L. Hu, S. Saito, L. Wei, *et al.*, “Avatar digitization from a single image for real-time rendering,” *ACM Trans. Graph.* **36**(4), 1–14 (2017).
6. W. Xie, Z. Kuang, and M. Wang, “SCIFI: 3D face reconstruction via smartphone screen lighting,” *Opt. Express* **29**(26), 43938–43952 (2021).
7. P. Han, Y. Cai, F. Liu, *et al.*, “Computational polarization 3D: New solution for monocular shape recovery in natural conditions,” *Opt. Lasers Eng.* **151**, 106925 (2022).
8. P. An, T. Ma, K. Yu, *et al.*, “Geometric calibration for LiDAR-camera system fusing 3D-2D and 3D-3D point correspondences,” *Opt. Express* **28**(2), 2122–2141 (2020).
9. W. Feng, T. Qu, J. Gao, *et al.*, “3D reconstruction of structured light fields based on point cloud adaptive repair for highly reflective surfaces,” *Appl. Opt.* **60**(24), 7086–7093 (2021).
10. W. Li, S. Shan, and H. Liu, “High-precision method of binocular camera calibration with a distortion model,” *Appl. Opt.* **56**(8), 2368–2377 (2017).
11. Y. Deng, J. Yang, S. Xu, *et al.*, “Accurate 3D Face Reconstruction With Weakly-Supervised Learning: From Single Image to Image Set,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (IEEE, (2019), pp. 285–295.
12. Y. Guo, J. Zhang, J. Cai, *et al.*, “CNN-Based Real-Time Dense Face Reconstruction with Inverse-Rendered Photo-Realistic Face Images,” *IEEE Trans. Pattern Anal. Mach. Intell.* **41**(6), 1294–1307 (2019).
13. A. Tewari, M. Zollhofer, P. Garrido, *et al.*, “Self-Supervised Multi-level Face Model Learning for Monocular Reconstruction at Over 250 Hz,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, 2018), pp. 2549–2559.
14. X. Li, F. Liu, P. Han, *et al.*, “Near-infrared monocular 3D computational polarization imaging of surfaces exhibiting nonuniform reflectance,” *Opt. Express* **29**(10), 15616 (2021).
15. X. Li, Z. Liu, Y. Cai, *et al.*, “Polarization 3D imaging technology: a review,” *Front. Phys.* **11**, 6 (2023).
16. P. Han, X. Li, F. Liu, *et al.*, “Accurate Passive 3D Polarization Face Reconstruction under Complex Conditions Assisted with Deep Learning,” *Photonics* **9**(12), 924 (2022).
17. L. B. Wolff, “Surface Orientation From Two Camera Stereo With Polarizers,” in *Optics, Illumination, and Image Sensing for Machine Vision IV* (SPIE, 1990), 1194, pp. 287–297.
18. L. B. Wolff, “Polarization vision: a new sensory approach to image understanding,” *Image Vis. Comput.* **15**(2), 81–93 (1997).
19. W. A. Shurcliff, *Polarized Light; Production and Use*. (Harvard University Press, 1962).
20. L. B. Wolff and T. E. Boulton, “Constraining object features using a polarization reflectance model,” *IEEE Trans. Pattern Anal. Machine Intell.* **13**(7), 635–657 (1991).
21. A. H. Mahmoud, M. T. El-Melegy, and A. A. Farag, “Direct method for shape recovery from polarization and shading,” in *2012 19th IEEE International Conference on Image Processing* (2012), pp. 1769–1772.
22. G. A. Atkinson and E. R. Hancock, “Recovery of surface orientation from diffuse polarization,” *IEEE Trans. on Image Process.* **15**(6), 1653–1664 (2006).
23. K. P. Gurton, A. J. Yuffa, and G. W. Videen, “Enhanced facial recognition for thermal imagery using polarimetric imaging,” *Opt. Lett.* **39**(13), 3857–3859 (2014).
24. A. Kadambi, V. Taamazyan, B. Shi, *et al.*, “Polarized 3D: High-Quality Depth Sensing with Polarization Cues,” in *2015 IEEE International Conference on Computer Vision (ICCV)* (2015), pp. 3370–3378.
25. Y. Yu, D. Zhu, and W. A. P. Smith, “Shape-from-Polarisation: A Nonlinear Least Squares Approach,” in *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)* (2017), pp. 2969–2976.
26. W. A. P. Smith, R. Ramamoorthi, and S. Tozza, “Height-from-Polarisation with Unknown Lighting or Albedo,” *IEEE Trans. Pattern Anal. Mach. Intell.* **41**(12), 2875–2888 (2019).
27. X. Xiao, B. Javidi, G. Saavedra, *et al.*, “Three-dimensional polarimetric computational integral imaging,” *Opt. Express* **20**(14), 15481 (2012).
28. K. Usmani, T. O’Connor, X. Shen, *et al.*, “Three-dimensional polarimetric integral imaging in photon-starved conditions: performance comparison between visible and long wave infrared imaging,” *Opt. Express* **28**(13), 19281 (2020).
29. S. Rajan, S. Wang, and R. Inkol, “Efficient Approximations for the Four-Quadrant Arctangent Function,” in *2006 Canadian Conference on Electrical and Computer Engineering* (2006), pp. 1043–1046.

30. D. Zhu and W. A. P. Smith, "Depth From a Polarisation + RGB Stereo Pair," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019), pp. 7578–7587.
31. R. T. Frankot and R. Chellappa, "A method for enforcing integrability in shape from shading algorithms," *IEEE Trans. Pattern Anal. Machine Intell.* **10**(4), 439–451 (1988).