PolarTact3D: Single-shot Tactile 3D Shape and Color Sensing with Polarization Imaging

Kai Garcia Mairi Yoshioka Huaijin Chen
Information and Computer Science Dept.
University of Hawaii, Manoa
Honolulu, HI 96822, USA

Tyler Ray Tianlu Wang
Mechanical Engineering Dept.
University of Hawaii, Manoa
Honolulu, HI 96822, USA

Frances Zhu
Mechanical Engineering Dept.
Colorado School of Mines
Golden, CO 80401, USA

https://polartact3d.github.io/

***Abstract*—Vision-based tactile sensors are essential for robotic applications such as grasping and physical interaction. We propose a low-cost, polarization-imaging-based tactile sensor that captures both shape and color information in a single shot. Unlike photometric-stereo-based solutions like GelSight, which require precise internal illumination and cannot reliably capture color due to the reflective coating required for surface reconstruction, our method leverages the Angle of Linear Polarization (AoLP) and Degree of Linear Polarization (DoLP) to encode 3D geometry. This approach enables robust shape reconstruction, even on transparent or specular targets. The sensor is constructed using commercial transparent polyethylene (PE) film and an off-the-shelf polarization camera, making it simple and inexpensive to build. We demonstrate the effectiveness of our design through real-world experiments on various contact surface scenarios.**

I. Introduction

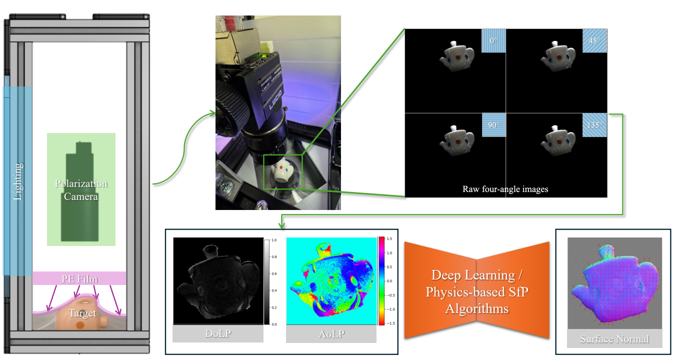

Fig. 1. **System Overview**. The proposed system consists of a polarization camera mounted inside an aluminum cage, facing downward. An elastic sheet of transparent, glossy polyethylene (PE) film is stretched across the bottom of the cage, and background illumination is provided from the sides. During measurement, the user firmly presses the cage against the target object, allowing the PE film to conform to the object's surface, serving as a dielectric coating layer to the object to enable the shape from polarization (SfP) algorithms to reconstruct the surface normal of the object, using either neural networks or physics-based methods.

Tactile 3D sensing plays a critical role in robotic manipulation, object recognition, and contact-rich interactions. A prominent example is the GelSight sensor [5], which uses photometric stereo to recover detailed surface geometry through elastomer deformation. Variations of this approach, such as in [12], have demonstrated impressive accuracy and grasping capability. However, GelSight-style sensors require precisely calibrated internal lighting and a flat, reflective interface for the photometric stereo principle it relies on, adding mechanical complexity, limiting robustness, and making simultaneous color sensing difficult. Recent work, DTact [6, 7], addresses some of these constraints by employing deformation-induced optical darkening for 3D shape modeling instead of photometric stereo, using a stack of semi-transparent elastomer layers and only external lighting. While this innovation reduces hardware complexity and increases adaptability, it still lacks the ability to recover surface color.

In this work, we take a first step toward a new class of vision-based tactile sensors by leveraging polarization imaging. By using a thin sheet of transparent glossy PE film serving as the interface between the object and a polarization camera, we exploit the angle-dependent reflection and polarization patterns to infer surface geometry. Our system enables simultaneous shape and color sensing in a single image without the need for calibrated internal lighting, and can generalize to non-planar contact surfaces. This approach simplifies the sensor construction while opening up new directions for low-cost, compact, high-resolution tactile sensing.

II. Shape from Polarizaion

Polarization imaging captures the orientation of light-wave oscillation, which provides geometric and material information for the surface [4, 2]. A typical polarization camera measures light intensity through four linear polarizing filters oriented at 0°, 45°, 90°, and 135°, captured in four images of $I_0, I_{45}, I_{90}, I_{135}$, allowing estimation of the polarization state at each pixel described by the Stokes parameters S_0, S_1, and S_2, which are computed as follows: $S_0 = I_0 + I_{90}$, $S_1 = I_0 - I_{90}$, and $S_2 = I_{45} - I_{135}$. The two key quantities of polarization can be calculated from the Stokes parameters:

- Degree of Linear Polarization (DoLP): $\rho = \sqrt{S_1^2 + S_2^2}/S_0$, which quantifies the proportion of light that is linearly polarized.
- Angle of Linear Polarization (AoLP): $\varphi = \frac{1}{2}\tan^{-1}(S_2/S_1)$, which represents the dominant orientation of the electric field vector.

As Fig. 2 shows, the measured DoLP and AoLP from the sensor encodes the surface normal of the target. For specular reflection, based on the dichromatic reflection model [1, 11, 10], the DoLP ρ_s is given by Eq.1,

$$\rho_s = \frac{2\sin^2\theta\cos\theta\sqrt{\eta^2 - \sin^2\theta}}{\eta^2 - \sin^2\theta - \eta^2\sin^2\theta + 2\sin^4\theta} \tag{1}$$

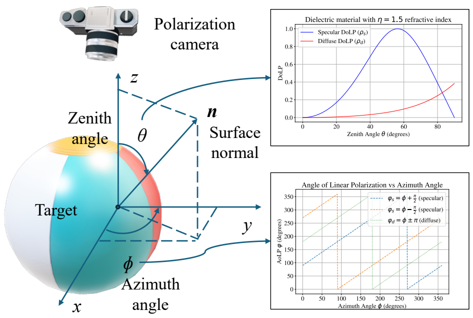

Fig. 2. **Surface normal from polarization**. A polarization camera can capture both the DoLP and the AoLP at each pixel location, which encode 3D surface geometry. For a dielectric material with known refractive index η, the measured DoLP is dependent on the zenith angle θ of the 3D surface normal vector using Eq. 1 and 3, while the azimuth angle ϕ maps directly to AoLP using Eq. 2 and 4. Accurate recovery of the surface normal requires resolving the ambiguities between specular and diffuse reflection and information about the material. In the proposed system, we apply a glossy polyethylene (PE) film over the surface, which has an known $\eta = 1.5$ and ensures dielectric polarization response and suppresses diffuse ambiguity, enabling shape inference across arbitrary underlying materials.

which depends on the zenith angle of the surface θ, given a refractive index η, with maximum DoLP occurring near the Brewster angle. The corresponding AoLP,

$$\varphi_s = \phi \pm \frac{\pi}{2} \tag{2}$$

is oriented perpendicular to the plane of reflection, and thus offset by $\pm\frac{\pi}{2}$ from the azimuth angle ϕ of the surface normal.

For diffuse reflection, derived from the Fresnel equation [13, 14, 1, 11], the DoLP ρ_d, expressed in Eq. 3,

$$\rho_d = \frac{(\eta - 1/\eta)^2 \sin^2\theta}{2 + 2\eta^2 - (\eta + 1/\eta)^2 \sin^2\theta + 4\cos\theta\sqrt{\eta^2 - \sin^2\theta}} \tag{3}$$

arises from subsurface scattering and internal reflection, with polarization strength governed by the zenith angle. The AoLP in this case,

$$\varphi_d = \phi \pm \pi \tag{4}$$

aligns with the plane of incidence and is offset by $\pm\pi$ from the azimuthal direction. As such, the surface normal can be inferred from polarization measurements. However, ambiguities remain, such as reflection type, ρ_s, and ϕ's $\pm\pi$ or $\pm\frac{\pi}{2}$ offsets, making exact angle recovery challenging without additional priors. In our system, a glossy PE film serves as the contact interface, reducing diffuse/specular ambiguity and has a known $\eta = 1.5$.

III. Proposed System

We constructed an aluminum-extrusion rig housing a Lucid PHX050S-QC polarization color camera that views an elastic, glossy PE film stretched across the bottom window as the tactile interface, operating in real-time at over 25 fps. A programmable LED ring provides uniform side illumination. During acquisition, the rig is pressed onto the object so the film conforms to the surface, acting as a thin dielectric coating for SfP. The captured frames are currently processed offline. The physics-based SfP can run in real time at over 30 fps, while the neural network-based SfP takes 1.5 seconds per frame on an NVIDIA RTX4090 GPU. We use the polarization camera to capture the raw mosaicked images with 0°, 45°, 90°, and 135° polarization angles in a single shot through the sensor's focal plane array division. We demosaic the raw image [9] and calculate the DoLP and AoLP images. For physics-based SfP, we build a look-up-table between ρ_s, φ_s and θ, ϕ, respectively using Eq. 3 and 4. For ambiguities, we ignore ρ_s larger than the Brewster angle and only maps $\varphi_s = -\frac{\pi}{2}$. For neural network SfP, we directly pass the DoLP and AoLP images to the SfPUEL network [8] to infer the surface normal.

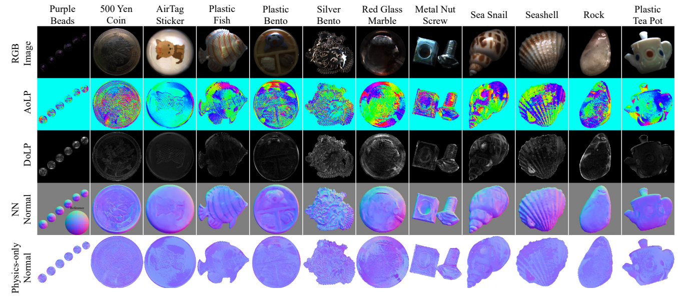

Fig. 3. **Results.** RGB, DoLP, AoLP, and normal map from NN and physics-only methods of various objects are provided. See Sec. III for discussion.

IV. Results

Figure 3 shows real-world shape and color results from our system across diverse materials—including conductors, dielectrics, and insulators—with shiny, diffuse, transparent, and translucent appearances. Both physics-based and neural network (NN)-based shape-from-polarization (SfP) methods are presented, along with a reference sphere for surface normal visualization. The physics-based approach suffers from ambiguities due to incomplete zenith look-up tables and $-\frac{\pi}{2}$ azimuth assumptions (See Sec. III). The NN method often yields more plausible surface normals with fine details, as seen in the AirTag sticker (note the sticker's thin layer and micro surface structure) and silver necklace. However, in low-DoLP regions, it may hallucinate geometry, likely relying on RGB cues. For example, the fish's colored strip appears incorrectly bulged. Transparent objects like glass remain challenging due to unmodeled inter-reflections.

V. Conclusion and Future Work

This work takes an initial step toward polarization-based tactile 3D and color sensing. Our prototype shows that meaningful shape information can be extracted through a conformal PE film using polarization cues. While promising, the system remains a proof of concept with several limitations. Results are currently qualitative; future work will include quantitative evaluations of surface normal accuracy. We also aim to miniaturize the hardware and incorporate more physics-based priors into SfP neural networks [11]. Directional lighting from a programmable LED array may further enhance inference. Additionally, stress sensing could be achieved via a polariscope configuration [3]. Mechanically, the thin PE film can "tent" over discontinuities, reducing accuracy. Future designs may adopt thicker, gel-like materials such as PDMS for improved surface conformity and polarization response.

References

[1] Gary A Atkinson and Edwin R Hancock. Recovery of surface orientation from diffuse polarization. *IEEE transactions on image processing*, 15(6):1653–1664, 2006.

[2] Seung-Hwan Baek, Ramesh Raskar, Jinwei Ye, Akshat Dave, Achuta Kadambi, and Huaijin Chen. Polarization-based visual computing. In *ACM SIGGRAPH 2023 Courses*, pages 1–1. 2023.

[3] Akshat Dave, Tianyi Zhang, Aaron Young, Ramesh Raskar, Wolfgang Heidrich, and Ashok Veeraraghavan. Nest: Neural stress tensor tomography by leveraging 3d photoelasticity. *ACM Transactions on Graphics*, 44(2): 1–18, 2025. doi: 10.1145/3723873.

[4] Achuta Kadambi, Vage Taamazyan, Boxin Shi, and Ramesh Raskar. Polarized 3d: High-quality depth sensing with polarization cues. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3370–3378, 2015.

[5] Rui Li and Edward H Adelson. Sensing and recognizing surface textures using a gelsight sensor. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1241–1247, 2013.

[6] Changyi Lin, Ziqi Lin, Shaoxiong Wang, and Huazhe Xu. Dtact: A vision-based tactile sensor that measures high-resolution 3d geometry directly from darkness. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 10359–10366. IEEE, 2023.

[7] Changyi Lin, Han Zhang, Jikai Xu, Lei Wu, and Huazhe Xu. 9dtact: A compact vision-based tactile sensor for accurate 3d shape reconstruction and generalizable 6d force estimation. *arXiv preprint arXiv:2308.14277*, 2023.

[8] Youwei Lyu, Heng Guo, Kailong Zhang, Si Li, and Boxin Shi. SfPUEL: Shape from polarization under unknown environment light. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024.

[9] Ryota Maeda. Polanalyser: Polarization image analysis tool, 2019. URL https://github.com/elerac/polanalyser.

[10] Steven A Shafer. Using color to separate reflection components. *Color Research & Application*, 10(4):210–218, 1985.

[11] Mingqi Shao, Chongkun Xia, Zhendong Yang, Junnan Huang, and Xueqian Wang. Transparent shape from a single view polarization image. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9277–9286, 2023.

[12] Shaoxiong Wang, Yu She, Branden Romero, and Edward Adelson. Gelsight wedge: Measuring high-resolution 3d contact geometry with a compact robot finger. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6468–6475. IEEE, 2021.

[13] Lawrence B. Wolff and Terrance E. Boult. Constraining object features using a polarization reflectance model. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 13(07):635–657, 1991.

[14] Lawrence B Wolff, Shree K Nayar, and Michael Oren. Improved diffuse reflection models for computer vision. *International Journal of Computer Vision*, 30:55–71, 1998.