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

Kai Garcia [1] Mairi Yoshioka [1] Huaijin Chen [1] Tyler Ray [2] Tianlu Wang [2] Frances Zhu [3]

[1]Department of Information and Computer Science [2]Department of Mechanical Engineering, University of Hawaii at Mānoa

[3]Department of Mechanical Engineering, Colorado School of Mines

Overview

Most tactile sensors like GelSight rely on internal LEDs and reflective gels, which **block colour** and complicate hardware. **PolarTact3D** captures both *shape & colour* in a single shot (25 fps), using only a glossy PE film and a polarization camera.

Hardware Prototype

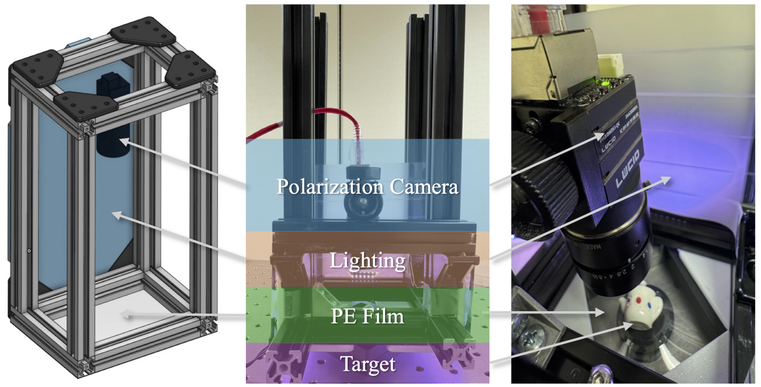

figure: Aluminium cage, Lucid PHX050S-QC polarisation camera, side LEDs, PE film interface.

Data Collection

- 4-angle polarisation data captured at 25 fps
- SfPUEL: trained on synthetic dome dataset (Lyu et al.)
- Evaluation: coins, jewelry, glass, plastic toys, and food models

Why Polarization?

- Encodes surface orientation through DoLP and AoLP
- Works with passive lighting — no structured light needed
- Sensitive to subtle surface textures invisible to RGB
- Allows color data collection and works on transparent materials

Shape From Polarisation

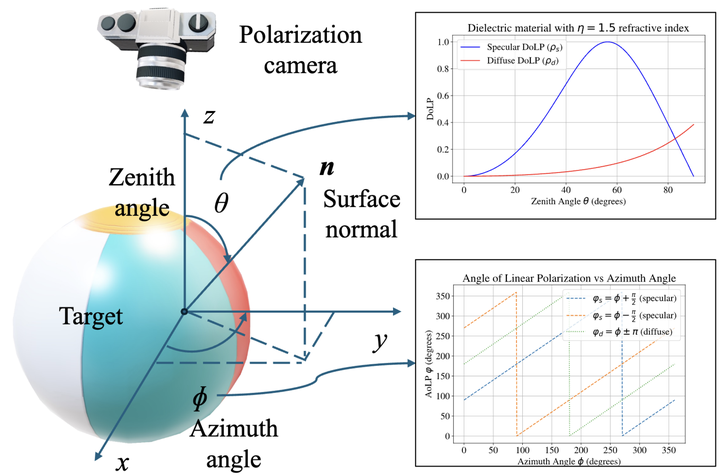

figure: DoLP vs zenith θ (η = 1.5) and AoLP ambiguities for specular vs diffuse reflection. The PE film enforces a known η and suppresses diffuse ambiguity.

System Pipeline

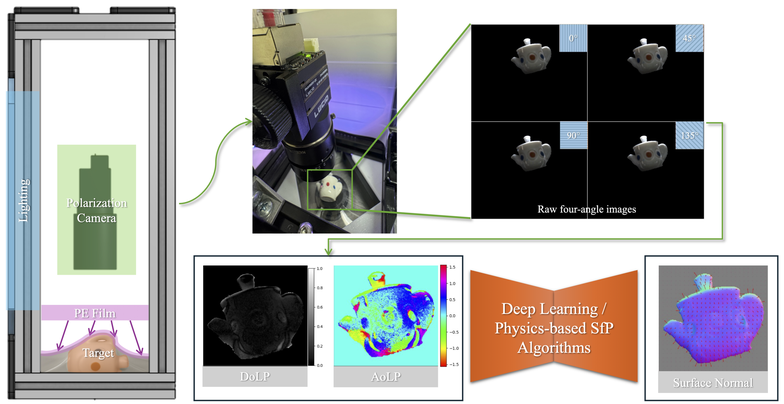

figure: From raw 4-angle mosaic → DoLP/AoLP → SfPUEL NN or physics LUT → surface normals.

Qualitative Results

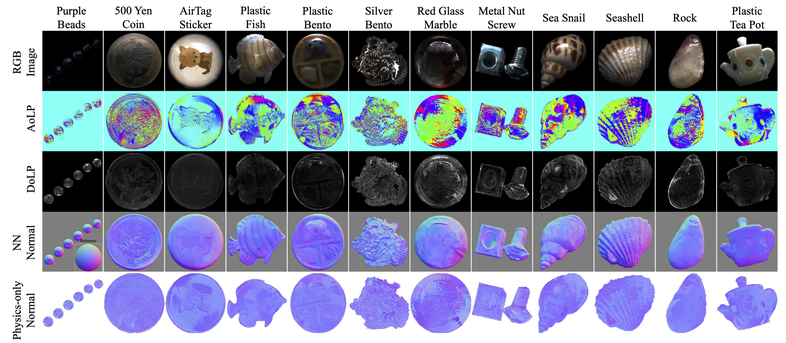

figure: Top→bottom: unpolarised RGB, AoLP, DoLP, SfPUEL normals, Physics-based normals. Diverse targets: beads, 500-yen coin, dog tag, fish, lunch toy, necklace, screws, teapot.

Limitations & Future Work

Current limitations

- No quantitative surface-normal RMSE yet (qualitative only).
- NN may *hallucinate* geometry in low-DoLP zones.
- Transparent objects with strong inter-reflections remain hard.
- Thin PE film can "tent" over sharp edges/discontinuities.
- SfPUEL pipeline ≈ 1.5 s / frame (GPU) — not real-time.

Ongoing work

- Benchmark normals vs. GT spheres; publish RMSE table.
- Thicker PDMS layer for better conformity and stress sensing.
- Directional LED array + physics priors inside the NN.
- Miniaturised board-camera version; on-device inference.

Conclusion

- **One-shot, two outputs** — high-resolution *shape & colour* in a single polarization frame.
- **Dual inference paths**
  - Physics-based LUT
  - SfPUEL U-Net++
- **Works on anything** — shiny, diffuse, translucent, even transparent targets.

Project Website Github