Scene depths from a two-polarization metalens

Brand, Matthew; Kuang, Zeyu

TR2023-105 September 01, 2023

Abstract

We develop a regularized depth-from-polarization formulation that works with as little as two distinctly polarized images, and optimize a polarization-sorting metalens for accurate single-exposure depth estimates in this framework.

Optica Imaging Congress / Flat Optics 2023

Scene depths from a two-polarization metalens

Matt Brand[1,*] and Zeyu Kuang[1,2]

[1] *Mitsubishi Electric Research Labs (MERL), 201 Broadway, Cambridge MA, USA*
[2] *Dept. Applied Physics, 15 Prospect St., Yale University, New Haven CT, USA*

[*] *brand@merl.com*

Abstract: We develop a regularized depth-from-polarization formulation that works with as little as two distinctly polarized images, and optimize a polarization-sorting metalens for accurate single-exposure depth estimates in this framework.

Polarized images of a scene contain information about surface orientations, material properties, and scene illumination. There is a well-developed literature on computing surface normals from multiple, distinctly polarized images, usually in multiplexed combination with multiple, distinct illumination sources. The key result is that determining a surface normal requires 3-4 distinctly polarized images *plus* information about the lighting direction [5]. Normals can then be integrated to get an estimate of the surface shape. Due to noise sensitivity, the general practice is to collect several more images to better constrain the estimates. We introduce a regularization that enables surface depth reconstruction directly from as little as two polarizations obtained in a single exposure through a polarization-sorting metalens. To improve peformance, we make the design and propagation simulation of such metalenses continuously differentiable so that the metalens can be tuned to support better depth estimates.

Design and optimization of a polarization-sorting metalens: The unit cell decomposition (UCD) decomposes a metasurface into a grid of subwavelength-sized atoms, each of which can be chosen independently to provide a local phase delay in the near field. Observing that the phase delay of silicon nitride nanopillars depends on the polarization direction of the incident wavefront, Miyata *et al.* [2] made a table of the phase delays in horizontally and vertically polarized light provided by different pillar geometries, and used this to piecewise design a metalens that focuses $0^\circ, 45^\circ, 90^\circ, 135^\circ-$polarized images to spatially separated focii on the focal plane. We take a similar approach to design a $0^\circ \,\&\, 90^\circ-$polarized focuser, but replace the table with a differentiable function approximator (e.g., bicubic regression) and also differentiate through the Rayleigh propagator so that the nanopillars' design parameters can be tuned to optimize any measure of performance at or beyond the focal plane. For example, by maximizing the intensity of polarized light focused in a small region around the appropriate focal point, we are able to improve focal efficiency by up to 5%, as determined from rigorous coupled-wave analysis (RCWA).

Depth from polarization: When light interacts with a surface, the perpendicularly polarized components are more likely to be absorbed or refracted, leaving the reflected light more parallel polarized. The intensity of reflected light measured by an observer depends on the illumination, illumination direction, surface orientation, surface refractive index, observer direction, and polarization. As it will turn out, some of these unknowns can be cancelled out by taking ratios of intensities measured with different polarizations, and the remaining unknowns can be handled by solving for a smooth surface whose slopes are consistent with spatial variations in these ratios.

We start with a surface point p whose normal, in spherical coordinates, is $\phi(p)$ off and $\theta(p)$ around the camera axis, i.e., a normal with $\phi = 0$ points directly at the camera. The surface at p has refractive index $\eta(p)$ and total diffuse albedo $\rho(p)$. For light reflecting off a smooth dielectric surface, we use the Wolff model [4] for the radiance of reflected light after two Fresnel transmissions (in and out):

$$L_r = \rho L \cdot (1 - F\{\psi, \eta(p)\})(\cos\psi)\left(1 - F\left\{\sin^{-1}\left[\sin\varepsilon/\eta(p)\right], \eta(p)^{-1}\right\}\right) d\omega \tag{1}$$

Here light is incident on p with radiance L at incidence angleÂ ψ through a small solid angle $d\omega$, and the reflected light is emitted at an emittance angle ε. The term $F(,)$ is the standard Frensel reflection coefficient.

The intensity of light received from the surface is $I_r = dA\frac{\cos\varepsilon_r}{r^2}L_r$, whereÂ dA is the surface area, θ_r is the emittance angle, andÂ r is the surface-to-receiver distance. The camera measures the transmitted radiance sinusoid [3]:

$$I(\ell, \upsilon, p) = (I_{\max}(\ell, p) + I_{\min}(\ell, p) + I_{\max}(\ell, p) - I_{\min}(\ell, p)\cos(2\nu - 2\phi(p)))/2$$

where $I(\ell, \upsilon, p)$ is the intensity of light received from surface point p illuminated by a light along direction ℓ and viewed at with a polarizer at angle ν. $I_{\max}(\ell, p), I_{\min}(\ell, p)$ are, respectively, the maximum and minimum radiances over all possible polarizer angles ν; their sum is the total intensity I_r.

The degree of polarization (DoP) [1] observable by the camera is $d(\theta(p), \eta(p)) \doteq$

$$\frac{I_{\max}(\ell, p) - I_{\min}(\ell, p)}{I_{\max}(\ell, p) + I_{\min}(\ell, p)} = \frac{(\eta(p) - \eta(p)^{-1})^2 sin^2\theta(p)}{2 + 2\eta(p)^2 - (\eta(p) + \eta(p)^{-1})^2 \sin^2\theta(p) + 4\cos\theta(p)\sqrt{\eta(p)^2 - \sin^2\theta(p)}}$$

where the expansion is obtained by plugging in appropriate Fresnel and Snel formulas [3]. The DoP is useful for rewriting the sinusoid as

$$I(\ell, \nu, p) = I_r(1 + d(\theta(p), \eta(p))\cos(2\nu - 2\phi(p)))/2$$

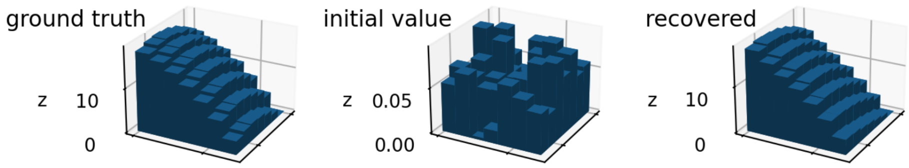

Fig. 1. Reconstruction of a doubly-curved surface from a 2-polarization single exposure.

from which it follows that the ratio of intensities at p seen through a polarizer at two angles ν_1, ν_2 is

$$\frac{I(\ell,\nu_1,p)}{I(\ell,\nu_2,p)} = \frac{1+d(\theta(p),\eta(p))\cos(2\nu_1-2\phi(p))}{1+d(\theta(p),\eta(p))\cos(2\nu_2-2\phi(p))}$$

Note that the unknown light source is cancelled out. The metalens camera gives us the left side measurements and ν_1, ν_2 as design data; the right side has three scene unknowns $\theta(p), \phi(p), \eta(p)$. Prior calculations on the amount of information needed to compute surface normals indicate that 4 polarization images are required to support stable estimates [6], or 2 polarization images taken in each of 3 lighting conditions [3]. To proceed with less, we assume that the observed surface is mostly smooth, such that most of the imaged pixels do not view depth discontinuities such as occlusion edges, and the refractive index varies slowly if at all. This is a good assumption for solid objects, albeit a poor one for highly porous volumes such as steel wool or shrubbery.

We now make a change of variables: Instead of characterizing surface orientation in terms of the normal direction $\theta(p), \phi(p)$, we rewrite the above two equations in terms of the surface *slope* as determined from the world coordinates of p's neighbors, of which only the depth values (collected in a vector $\mathbf{z}$) are unknown. The above ratio can be rewritten in the form $I_{\nu_1,i}/I_{\nu_2,i} = \hat{I}_{\nu_1,i}(\mathbf{z})/\hat{I}_{\nu_2,i}(\mathbf{z})$ where the left hand is our ratio of measured intensities for the i^{th} point and the right hand are predicted intensities for that point given a vector of depth estimates for all points. Letting $\nu_1 = x$ to indicate horizontal polarization and similarly $\nu_2 = y$ to indicate vertical, we rearrange this equality into a cross-ratio $I_{x,i}\hat{I}_{y,i}(\mathbf{z}) - I_{y,i}\hat{I}_{x,i}(\mathbf{z}) = 0$, yielding our ultimate optimization objective

$$\min_{\mathbf{z}=\{z_1,\cdots,z_N\}} \sum_i \left|I_{y,i}\hat{I}_{x,i}(\mathbf{z}) - I_{x,i}\hat{I}_{y,i}(\mathbf{z})\right|^2 + \lambda \sum_{i,j|i\in\mathcal{N}(i)} \left(1 - \langle \hat{n}_i(\mathbf{z}), \hat{n}_j(\mathbf{z})\rangle\right)^2 \tag{2}$$

Here the second term penalizes overfitting to noisy measurements by penalizing small cosines between implied normals $\hat{n}_i(\mathbf{z})$ of adjacent points. This objective can be differentated w.r.t. $\mathbf{z}$ and solved in a generic optimizer.

The remaining unknown in this problem is the refractive indices at the points. Assuming patchwise smoothness, these can also be estimated in an interleaved optimization using a total variation regularization.

Metals: The case of reflection off a metal surface follows a similar logic, but the intensity of the s and p polarization in the reflected light are now given by the total reflected intensity (1) times the portion of the s (or p) wave in the reflected light. That is determined by the last step of the transmission from dielectric to air, where the portion of the s-polarized light isÂ $|t_s|^2/(|t_s|^2+|t_p|^2)$ and the portion of the p-polarized light is $|t_p|^2/(|t_s|^2+|t_p|^2)$, where $t_s = \frac{2\cos\gamma}{\cos\gamma+\eta\cos\phi}, \gamma = \sin^{-1}(\eta^{-1}\sin\phi)$ and $t_p = \frac{2\cos\gamma}{\eta\cos\gamma+\cos\phi}$ are Frensel transmission coefficients of light refracting out of a surface at p whose normal makes angle ϕ to the sightline from the camera to p.

Example: To illustrate, we assume a diffuse unpolarized light source such as a cloudy sky, and a doubly curved surface made of copper, which has refractive index ratio (air-to-copper) of $\eta = 1.25+2.39i$ at 450nm. A doubly curved surface is positioned $100\mu m$ in front of the metalens and reflected light is propagated from $9\times 9 = 81$ distinct patches on the surface through the metasurface to a simulated CCD sensor placed at the focal plane. Intensity measurements are integrated at the corresponding CCD sensor wells and contaminated with 2% i.i.d. Gaussian noise. The objective (2) is then minimized by a generic optimizer, correctly recovering the scene geometry. The optimization result remains almost constant as regularizer weight λ is varied over 5 orders of magnitude, indicating that the smoothness constraint is making the problem well-posed without distorting the optimum. A slight curvature deficiency in the reconstruction indicates that the sample points should be closer so that finite differencing of their coordinates provides better approximations to the local surface slope.

References

1. E. Hecht. *Optics*. Addison Wesley, 4 edition, 2001.
2. M. Miyata, M. Nakajima, and T. Hashimoto. Compound-eye metasurface optics enabling a high-sensitivity, ultra-thin polarization camera. *Optics Express*, 28:9996–10014, 2020.
3. T. T. Ngo, H. Nagahara, and R.-I. Taniguchi. Surface normals and light directions from shading and polarization. *IEEE PAMI*, 2021.
4. L. B. Wolff. Diffuse-reflectance model for smooth dielectric surfaces. *JoSA A*, 11(11):2956–2968, 1994.
5. T. Wolff and L. Boult. Constraining object features using a polarization reflectance model. *IEEE PAMI*, 13(7):635–637, 1991.
6. L. Zhang and E. Hancock. Robust estimation of shape and polarisation using blind source separation. *Pattern Recognition Letters*, 34(8):856–862, 2013.