Supplementary material to: Height-from-Polarisation with Unknown Lighting or Albedo

1 Azimuth-from-boundary convexity prior

For objects with a foreground mask, we can obtain a coarse estimate of the azimuth angle of the surface normal at each pixel by assuming the object is globally convex. Under this assumption, the surface normal at the object boundary can be assumed to lie in the image plane, pointing outwards and perpendicular to the boundary tangent. Via an appropriate propagation scheme, these vectors can be propagated into the interior of the object, providing an azimuth estimate for all pixels.

We begin by computing 2D vectors that are normal to the boundary. We do this by tracing the boundary and using forward finite differences. Next, we propagate into the interior. There are two obvious ways to do this and we see no difference in performance between the approaches. The first method erodes the foreground mask and then compute normals to this eroded boundary in the same way as for the original boundary. This process is repeated until all pixels have been assigned a vector. The second approach computes the direction to the closest boundary pixel for each non-boundary pixel and uses this direction as the boundary vector. This can be averaged over the K closest boundary pixels for improved robustness.

The boundary normals estimated by this process are highly discretised. This is because there are only 8 possible directions for the outward facing normal as the boundary is traced. For this reason, we smooth the vector field estimated as described above. The precise details of the smoothing process do not affect results but we use an 11 by 11 window approximating a Gaussian kernel with a standard deviation of 2. Finally, the 2D vectors are converted to azimuth angles lying in the interval $[0, 2\pi)$. An example result is shown for the Stanford bunny in Figure 1 using the second method for propagation into the interior.

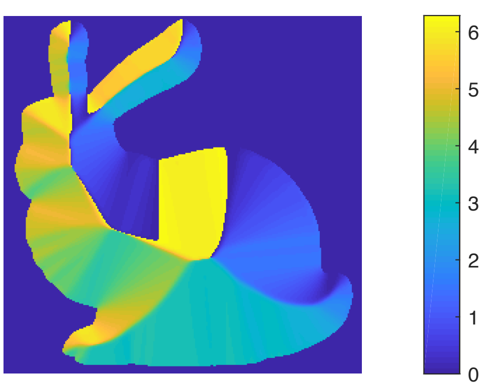

Fig. 1: Azimuth angle estimated from the foreground mask of the Stanford bunny under an assumption of global convexity.

2 Note on 2^K ambiguity

In the paper, we discuss the combinatorial set up of the lighting estimation problem. Since the polarisation normal at each pixel is subject to a binary ambiguity, there are in total 2^K possible disambiguations. Hence, we could build a linear system of equations for each one of these:

$$\begin{aligned}
\begin{bmatrix}\bar{\mathbf{n}}(\mathbf{u}_1) & \bar{\mathbf{n}}(\mathbf{u}_2) & \dots & \bar{\mathbf{n}}(\mathbf{u}_K)\end{bmatrix}^T \mathbf{s} &= \mathbf{i}, \\
\begin{bmatrix}\mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_1) & \bar{\mathbf{n}}(\mathbf{u}_2) & \dots & \bar{\mathbf{n}}(\mathbf{u}_K)\end{bmatrix}^T \mathbf{s} &= \mathbf{i}, \\
\begin{bmatrix}\bar{\mathbf{n}}(\mathbf{u}_1) & \mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_2) & \dots & \bar{\mathbf{n}}(\mathbf{u}_K)\end{bmatrix}^T \mathbf{s} &= \mathbf{i}, \\
&\vdots \\
\begin{bmatrix}\mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_1) & \mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_2) & \dots & \mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_K)\end{bmatrix}^T \mathbf{s} &= \mathbf{i}.
\end{aligned}$$

This is the principle behind Algorithms 1 and 2. In the case of Algorithm 1, $K = 4$ (the minimal case) and zero noise is assumed. In the case of Algorithm 2, $K \geq 4$, all pixels are used and the solution is least squares optimal. However, in both cases a subtle optimisation could be performed. The 2^K disambiguations can be divided into two halves such that each disambiguation in the first half differs from a disambiguation in the second half by a global transformation (i.e. all normals are transformed by $\mathbf{T}$). Concretely, suppose that the permutations are stored in $\mathbf{P} = \text{binaryStrings}(K)$ such that $P_{i,j}$ is the jth digit of the ith string. Then, for any string i there exists a string k such that $\forall j \in \{1, \dots, K\}, P_{i,j} = \neg P_{k,j}$. So, if the ith disambiguated normals are stored in $\mathbf{N}_i \in \mathbb{R}^{K\times 3}$ and the disambiguation for which all elements in the binary string were negated are stored in $\mathbf{N}_k \in \mathbb{R}^{K\times 3}$, then $\mathbf{N}_i = \mathbf{N}_k\mathbf{T}$. Moreover, the residuals of the two systems are equal:

$$\mathbf{N}_i\mathbf{N}_i^+\mathbf{i} - \mathbf{i} = \mathbf{N}_k\mathbf{N}_k^+\mathbf{i} - \mathbf{i}.$$

For this reason, there is no need to consider both halves of the 2^K disambiguations. Instead, we can consider only half, i.e. 2^{K-1}. We are still guaranteed to find the least squares optimal solution, which if globally transformed will yield the other optimal solution. For example, we might

θ_l	$\sigma = 0\%$	$\sigma = 0.5\%$	$\sigma = 1\%$	$\sigma = 2\%$
15°	0.62°	0.62°	0.64°	0.82°
30°	0.93°	1.03°	2.06°	4.61°
60°	1.83°	8.14°	14.2°	23.3°

TABLE 1: Quantitative point light source direction estimation results on synthetic data with varying noise. Errors are angular error.

arbitrarily decide to consider the half in which the first disambiguation is fixed to $\mathbf{N}_1 = \bar{\mathbf{n}}(\mathbf{u}_1)$. For the minimal case, this means that we only need to consider 8 disambiguations. While providing a practical optimisation, there is no asymptotic gain since $2^K = \Theta(2^{K-1})$.

3 Justification of synthetic noise level

In our experiments on synthetic data we vary additive Gaussian noise from $\sigma = 0.5\%$ to $\sigma = 2\%$. We now show that this is a reasonable approximation of the noise we can expect with our imaging setup. According to DxOMark[1], the Canon EOS-1DX has an SNR of 32.3dB at 15% grayscale intensity (this corresponds to a typical diffuse brightness in our captured data) when gain is set to ISO 800 (we used ISO 400 or 800 in our experiments). For the signal in our synthetic data ($\sigma_{\text{signal}} \approx 0.2$), this SNR translates to a noise std of 0.5%:

$$\sigma_{\text{noise}} = \sqrt{\frac{\sigma^2_{\text{signal}}}{10^{\frac{32.3\text{dB}}{10}}}} \approx= 0.5\%. \tag{1}$$

4 Additional experimental results

Here we provide additional experimental results. Specifically, we show the effect of varying the level of noise in the synthetic data experiments and perform an ablation study to examine the influence of the priors.

4.1 Light source estimation

In Table 1 we show light source estimation results for varying levels of noise and light source direction. Across all noise settings, the error is less than one degree for light sources 15° from the viewing direction. This only increases significantly in the presence of noise and when the light source is 60° from the viewing direction.

4.2 Surface height estimation

In Table 2 we show quantitative results for surface height estimation using the proposed method with known light source direction. In contrast to the results in the paper, here we vary the noise level over $\sigma \in \{0\%, 0.5\%, 1\%, 2\%\}$. Note that even with with $\sigma = 0\%$ there is still quantisation and saturation noise present. Moreover, the synthetic data uses a physically-based reflectance model whereas our shape recovery assumes Lambertian diffuse reflectance and Blinn-Phong specular reflectance. So, there is systematic noise in this model mismatch. Nevertheless, we show that we can obtain stable performance over a range of noise settings.

1. https://www.dxomark.com/Cameras/Canon/EOS-1Dx---Measurements

θ_l	$\sigma = 0\%$		$\sigma = 0.5\%$		$\sigma = 1\%$		$\sigma = 2\%$	
	Height (pixels)	Normal (degrees)	Height (pixels)	Normal (degrees)	Height (pixels)	Normal (degrees)	Height (pixels)	Normal (degrees)
15°	7.00	5.20	10.9	8.50	30.9	15.2	34.1	25.7
30°	6.91	4.89	9.80	6.86	16.1	12.3	25.3	20.4
60°	6.92	4.91	9.66	6.88	17.6	12.9	24.8	19.7

TABLE 2: Surface recovery results with varying noise.

Priors	$\sigma = 0\%$	$\sigma = 0.5\%$	$\sigma = 1\%$	$\sigma = 2\%$
Both	4.97	7.41	10.6	17.8
Smoothness only	5.81	8.07	11.5	19.3
Convexity only	7.21	10.1	14.2	23.9
None	7.53	10.7	15.4	25.0

TABLE 3: Results of an ablation study. We show mean angular error of the estimated surface normals, averaged over all light source directions.

4.3 Ablation study

We now show results of an ablation study where we examine the influence of the priors described in Sections 4.7 and 4.8. We evaluate performance using: 1. both priors, 2. Laplacian smoothness only, 3. convexity only, 4. neither. Results are shown in Table 3. For this experiment we use only uniform albedo, known light source direction and average over all light source directions. There is a roughly 30% gain by using both priors with the smoothness prior contributing significantly more than the boundary prior. However, it is worth noting that on real data the boundary prior does make a strong qualitative improvement by avoiding convex/concave errors near the boundary when the polarisation measurements are unreliable.