Practical Shape Recovery from Near-light and Near-camera Photometric Stereo

Xiang Huang, Oliver Cossairt and Marc Walton

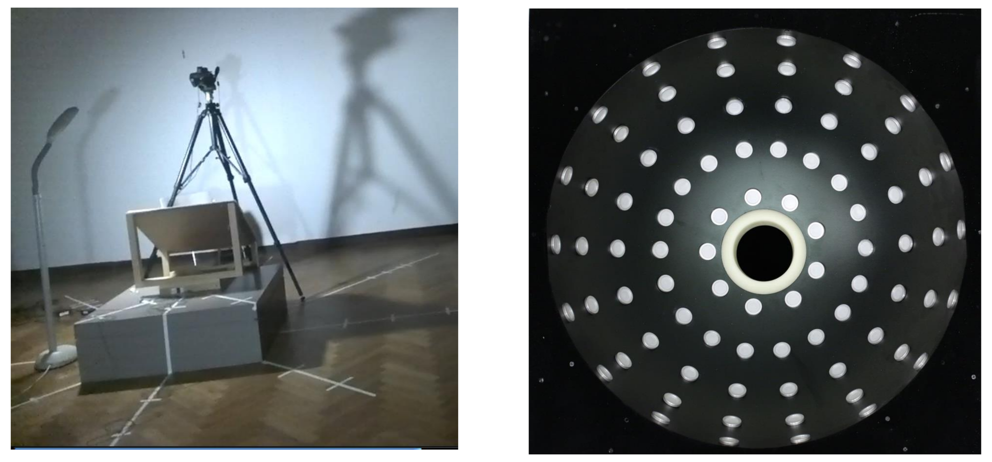
xianghuang@gmail.com

Department of Electrical Engineering and Computer Science, Northwestern University

Key Points

- *Near-light model for image relighting and 3D shape recovery;*
- *3D light position estimation using new optimization method with flat surface;*
- *Better image relighting vs Polynomial Texture Mapping: non-uniform spot light effect eliminated;*
- *Better surface normal and shape vs Reflectance Transformation Imaging: "potato-chip" shape error eliminated;*

Setup: Capture Images



(a) Hand-held flash light (b) Dome mounted with 81 LED lights

Figure: Two capture devices with fixed camera position and various lighting.

Image relighting: synthesis images with new light condition, by polynomial regression etc.

Photometric stereo: estimate surface normal from intensities, by Lambertian cosine law.

Both require light direction or position!

Image Intensity Model

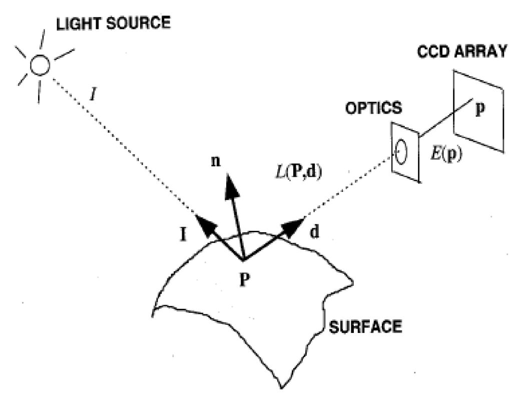


Figure: Light transport from the point light source, reflected by the object surface, and sensed by the camera.

Distant Light: $I_p = \hat{\boldsymbol{n}}_p \cdot \hat{\boldsymbol{l}} e a_p \eta_p = \hat{\boldsymbol{n}}_p^T \hat{\boldsymbol{l}} e a'_p$

Near Light: $I_p = \hat{\boldsymbol{n}}_p^T \frac{(\boldsymbol{l} - \boldsymbol{p})}{\|\boldsymbol{l} - \boldsymbol{p}\|^3} e a'_p$

$\hat{\boldsymbol{n}}_p$: normal direction $\quad$ $\boldsymbol{p}$: point position

$\hat{\boldsymbol{l}}$: light direction $\quad$ $\boldsymbol{l}$: light position

a_p : albedo $\quad$ η_p : vignetting factor

e : light power $\quad$ $a'_p = a_p \eta_p$: effective albedo

Optimization Solver

❖ *Energy Minimization for all pixels in K images:*

$$E(\boldsymbol{l_k}, e_k, \hat{\boldsymbol{n}}_p, \boldsymbol{p}, a'_p) = \sum_{p,k} \left(\frac{\hat{\boldsymbol{n}}_p^T (\boldsymbol{l_k} - \boldsymbol{p})}{\|\boldsymbol{l_k} - \boldsymbol{p}\|^3} e_k a'_p - I_{pk} \right)^2$$

Iteratively solve the following two sub-problems:

(I): Given the normal, shape and effective albedo, solve the lighting. Non-convex optimization due to the near-light term $\frac{(\boldsymbol{l} - \boldsymbol{p})}{\|\boldsymbol{l} - \boldsymbol{p}\|^3}$

(II): Given the relative lighting position, solve normal and effective albedo. This is a conventional problem can be solved by least-squares.

❖ ***New optimization method for step (II)***

- ***Minimize the following new objective function to find light positions. Note it eliminates the effective albedo variable.***

$$D = 1 - \frac{1}{NK} \sum_p \frac{\left(\sum_k I_{pk} g(\boldsymbol{l_k}, \boldsymbol{p}) e_k^{-1} \right)^2}{\sum_k \left(I_{pk} g(\boldsymbol{l_k}, \boldsymbol{p}) e_k^{-1} \right)^2}$$

$$g(\boldsymbol{l_k}, \boldsymbol{p}) \triangleq \frac{\|\boldsymbol{l_k} - \boldsymbol{p}\|^3}{\hat{\boldsymbol{n}}_p^T (\boldsymbol{l_k} - \boldsymbol{p})}$$

- ***Physical meaning:*** D is the variance of effective albedo a'_p normalized by its power, where $a'_p = I_{pk} g(\boldsymbol{l_k}, \boldsymbol{p}) e_k^{-1}$

- ***Solver***:
  - Compute gradient, use BFGS quasi-Newton method to find optimum light position
  - It is robust to initial value: simply initial lights at scene center
  - It is fast and converges in seconds: only need a small set of pixels

Image Relighting

❖ *Traditional far-light model: Polynomial Texture Mapping (PTM) or Hemi-spherical Harmonics (HSH) regression method*

$$I_p = c_0 l_x^2 + c_1 l_y^2 + c_2 l_x l_y + c_3 l_x + c_4 l_y + c_5 = \boldsymbol{h}(\hat{\boldsymbol{l}})^T \boldsymbol{c_p}$$

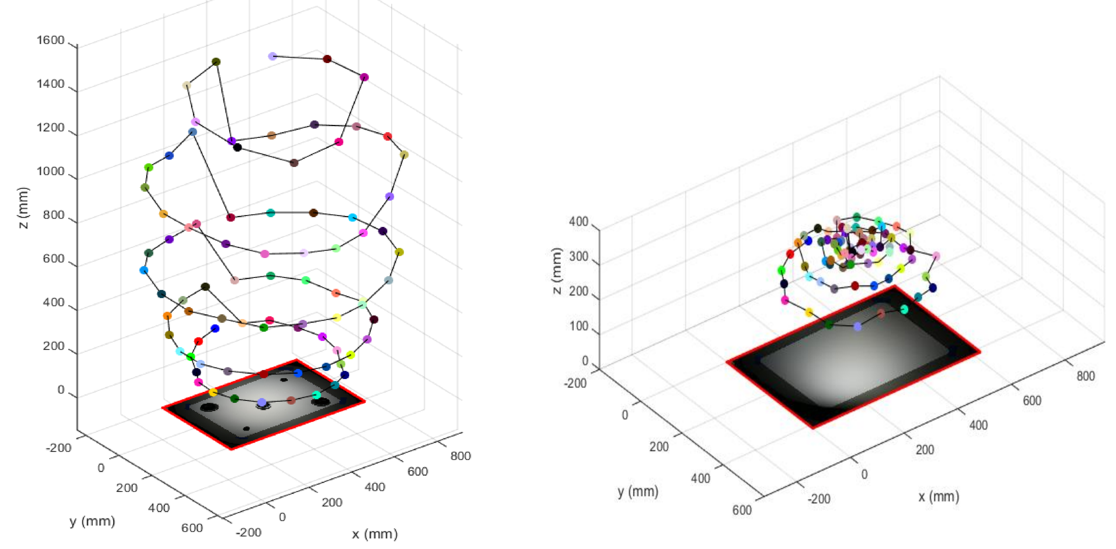
$$\begin{bmatrix} \boldsymbol{h}(\hat{\boldsymbol{l}}_1) \\ \boldsymbol{h}(\hat{\boldsymbol{l}}_2) \\ \vdots \\ \boldsymbol{h}(\hat{\boldsymbol{l}}_K) \end{bmatrix} \boldsymbol{c_p} = \begin{bmatrix} I_{p_1} \\ I_{p_2} \\ \vdots \\ I_{p_K} \end{bmatrix}$$

❖ *Our near-light model: modified PTM or HSH*

$$\begin{bmatrix} \boldsymbol{h}(\frac{(\boldsymbol{l_1} - \boldsymbol{p})}{\|\boldsymbol{l_1} - \boldsymbol{p}\|}) \\ \boldsymbol{h}(\frac{(\boldsymbol{l_2} - \boldsymbol{p})}{\|\boldsymbol{l_2} - \boldsymbol{p}\|}) \\ \vdots \\ \boldsymbol{h}(\frac{(\boldsymbol{l_K} - \boldsymbol{p})}{\|\boldsymbol{l_K} - \boldsymbol{p}\|}) \end{bmatrix} \boldsymbol{c_p} = \begin{bmatrix} I'_{p_1} \\ I'_{p_2} \\ \vdots \\ I'_{p_K} \end{bmatrix}$$

$$I'_{pk} = I_{pk} \frac{\|\boldsymbol{l_k} - \boldsymbol{p}\|^2}{e_k}$$

Results



(a)Light position from 5 mirror spheres (b) Our Light position from paper

Figure.: ***Light Position Estimation.*** A comparison of 3D light position estimation using triangulation from multiple mirror balls (left) and our method proposed (right). Left: the light positions of 81 dome lights obtained by least square error triangulation from five mirror spheres. Right: same lights estimated from just a piece of white matte printing paper. The light positions from the mirror balls are subject to large triangulation errors for lights near the top of the dome. Our technique, however, produces more accurate estimates of 3D light position.

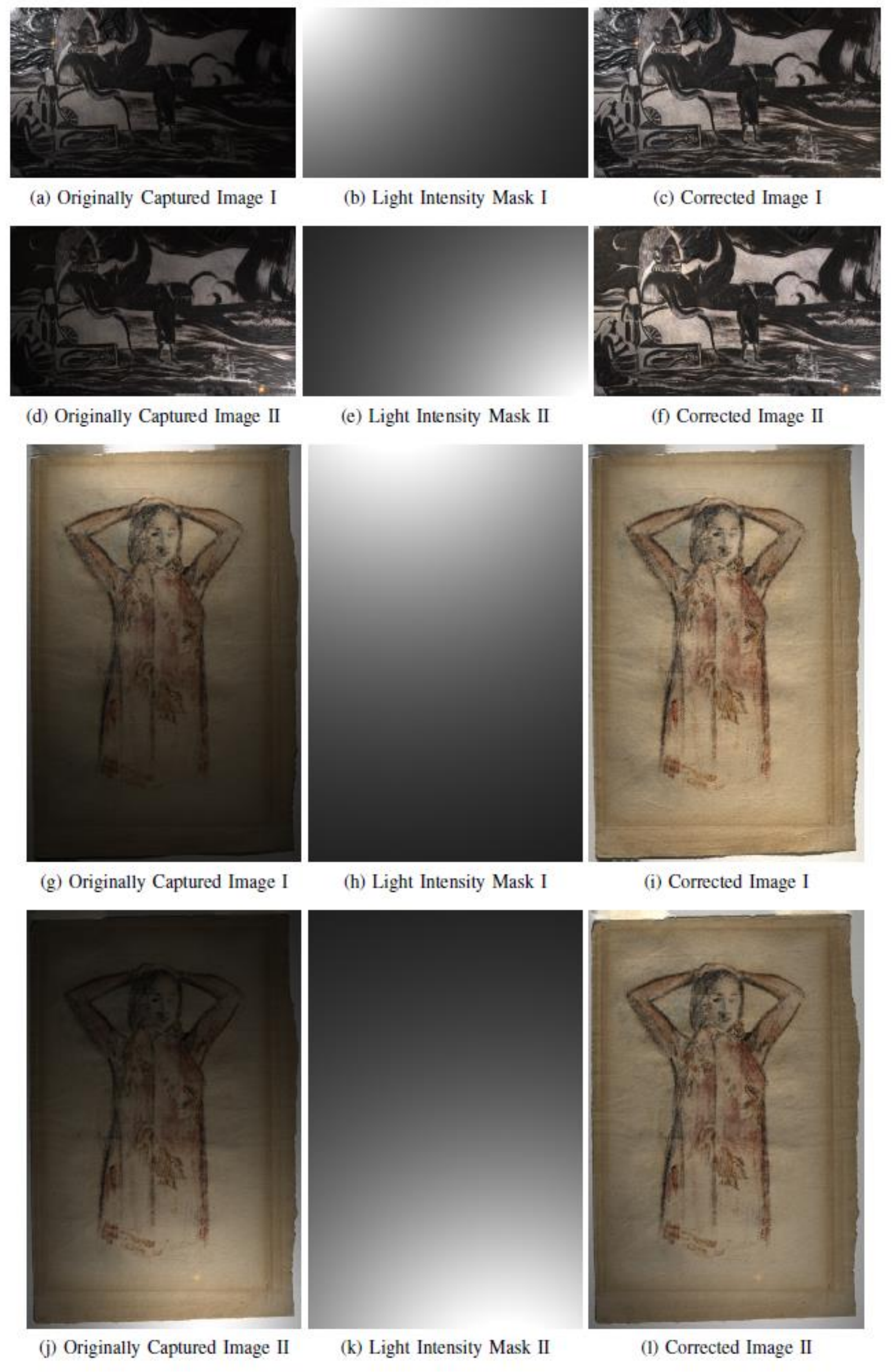


Figure: ***Image Relighting*** comparisons for two works by Paul Gauguin housed at the Art Institute of Chicago, a woodblock (top), and a transfer print (bottom). The raw captured images (left). After the near-light correction (right). We use the calibrated light position to compute the light attenuation due to the distance squared fall-off. The inverse of this attenuation mask (middle) is used to produce relit images with even illumination (right). The corrected images look uniformly lit and more visually pleasing.

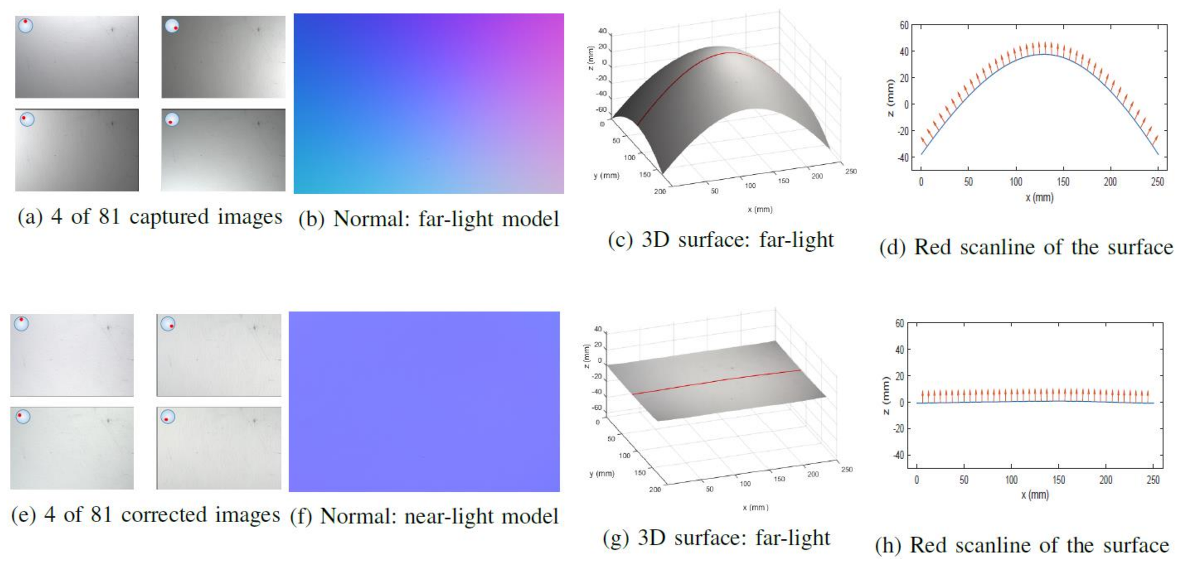


Figure.: ***Surface Normal and 3D shape Reconstruction.*** A comparison of far-light and near-light models. We use a flat matte paper as a ground truth test object. The first column shows the captured and corrected images. The position of the red dot in the circle approximates lighting direction. The second column shows the surface normal. The ground truth should be uniform, but the normal from the far-light model has a large error especially at the border. The third column shows the reconstructed 3d surface. The last column shows a horizontal scan line of the recovered depth map. The near-light model results are close to ground truth, while the far-light model results in a large potato chipped distortion.

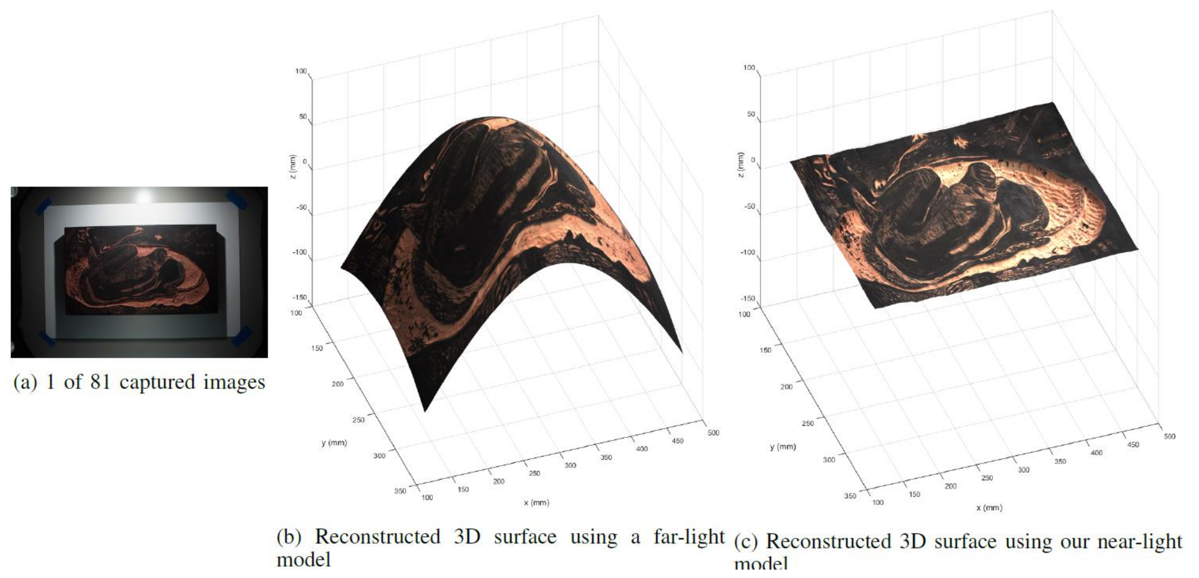


Figure.: ***3D Shape Reconstruction.*** A comparison of far-light and near-light models. The work of art shown is a woodblock produced by artist Paul Gauguin, housed at the Art Institute of Chicago. (a) One of the captured images showing the near-light illumination effect. (b) A 3D reconstruction without correcting for the near-light effect. The woodblock appears to be bent like a "potato chip" so surface details are difficult to resolve. (c) A 3D reconstruction using our method to correct for the near-light effect. The woodblock is now flat and details of the carving can be discerned from the reconstruction.

Conclusion

- *A novel automatic method to estimate 3D light locations*
- *Better image relighting and surface reconstruction for cultural heritage imaging applications compare with existing PTM etc.*

Future Work

- *Non-isotropic light source*
- *More sophisticated models to handle non-Lambertian effects*
- *What's the accuracy limit of photometric stereo?*

Can photometric stereo beat photogrammetry stereo?