Uncalibrated Near-Light Photometric Stereo

Thoma Papadhimitri
http://www.cvg.unibe.ch
Paolo Favaro
http://www.cvg.unibe.ch

Universität Bern
Institut für Informatik und Angewandte Mathematik
Bern, Switzerland

Abstract

In this work we solve the uncalibrated photometric stereo problem with lights placed near the scene. Although the devised model is more complex than its far-light counterpart, we show that under a global linear ambiguity the reconstruction is possible up to a rotation and scaling, which can be easily fixed. We also propose a solution for reconstructing the normal map, the albedo, the light positions and the light intensities of a scene given only a sequence of near-light images. This is done in an alternating minimization framework which first estimates both the normals and the albedo, and then the light positions and intensities. We validate our method on real world experiments and show that a near-light model leads to a significant improvement in the surface reconstruction compared to the classic distant illumination case.

1 Introduction

Photometric stereo (PS) [31] is a technique to accurately recover the normal map of a 3D scene from several pictures (at least three) taken from the same viewpoint and under different illumination conditions. When the light directions and intensities are known, photometric stereo can be solved as a linear system. When the illumination is not known, one needs to solve a much harder problem: uncalibrated photometric stereo. Typical assumptions are the Lambertian reflectance, orthographic projection, absence of shadows and interreflections and that the light sources are far away from the object. In particular, the last assumption allows to consider parallel illumination and, consequently, a simpler image formation model.

The distant light assumption is a reasonable approximation as long as the dimensions of the scene are much smaller than the distance of the light sources. However, this may not be the case in many practical scenarios such as endoscopy, cultural heritage, reconstruction of big indoor objects, underground and underwater navigation, or full human body 3D reconstruction. Recently, several hybrid methods [5, 34] that fuse photometric stereo and depth information provided by the Kinect [12] have been proposed. A near-light photometric stereo algorithm would be beneficial to this fusion especially for medium scale scenes (*e.g.*, the size of a room) when lights cannot be placed too far away from the objects. However, the close-lights calibration is not a trivial task [1, 6, 28, 29, 30].

Motivated by this fact, we introduce an uncalibrated near-light photometric stereo method where no prior information about light position and intensities is needed. We achieve this by first analyzing the reconstruction ambiguities and then by introducing an iterative technique to solve for the normals, reflectance and lights. We demonstrate the practical use and

accuracy of our algorithm with real world experiments and compare it with the state-of-art in uncalibrated distant light photometric stereo.

2 Prior Work

Near-Light Photometric Stereo: There are several shape from shading [14, 24, 32, 33] and calibrated photometric stereo [8, 10, 11, 16, 17, 19, 25] techniques that assume known illumination sources distributed near the object. In the early work of Clark [10] the depth map is estimated directly from intensities and their derivatives with respect to the light motion. For this reason a very precise controlled lighting setup is necessary which in practice is not always possible. In [16] a near-light photometric constraint is added to the multi-view stereo problem for 3D reconstruction. In [17] the depth is estimated via a non-linear equation. The uniqueness and convergence of the solution is guaranteed by assuming known the possible range of values the normals and depth can have. In a recent work of Ahmad et.al [8], the light locations are calculated using two chrome spheres. Later, the per-pixel light directions are obtained by subtracting from these locations the depth, previously estimated via the conventional distant light photometric stereo method. There are also techniques focusing on particular practical scenarios such as laparoscopy [11] or noise reduction [19]. Typically, the calibration is done by using at least two calibration spheres [22] and the 3D position of lights is calculated by triangulating at least two rays in space. This procedure is very sensitive to noise and measurement errors. Indeed, there is an extensive literature that considers these issues [1, 6, 28, 29, 30].

Unicalibrated Distant Light Photometric Stereo: In the distant light case, it is well known that when no information about the lights is available, the normals and lights can be obtained up to a 3×3 invertible matrix (nine-parameters ambiguity). This is done by computing a singular value decomposition of the data matrix and by forcing its rank to be 3 [15]. If the integrability constraint is imposed, the ambiguity reduces to the three-parameters generalized bas relief (GBR) ambiguity [7]. In order to impose the integrability constraint, a binary mask (typically generated with manual intervention) is necessary so that a smooth and continuous portion of the scene is available. In addition, the mask, which typically isolates a smaller and central portion of the scene, most of the times hides the reconstruction distortion (which is stronger towards the borders of the image) due to the distant light assumption. Several additional assumptions (on the geometry, the albedo and/or the lights) need to be imposed to the problem to fix the GBR ambiguity. This can be done by exploiting diffuse maxima [20], specular maxima [13], interreflections [9], entropy minimization [4, 23], chromaticity clustering [26] or perspective projection [21].

Unicalibrated Near-Light Photometric Stereo: In this work we focus for the first time on uncalibrated near-light photometric stereo. Only in [18] uncalibrated near-lights were considered. However, the method only recovers depth cues obtained from particular illumination configurations (lights moving on a line or plane), while in our algorithm we consider illuminants distributed arbitrarily in front of the object. We show that in the near-light case the ambiguities reduce to a rotation and scaling. Iteratively we update for the light positions, light intensities, normals, depth and albedo without making any stringent assumptions about scene geometry, reflectance or light distribution.

3 Near-Light Lambertian Model

The image formation model typically used for the near-light case under the Lambertian reflectance is (see [10, 14, 16, 32])

$$I_{pk} = \frac{\rho_p \mathbf{N}_p^T (\mathbf{L}_k - \mathbf{X}_p)}{\|\mathbf{L}_k - \mathbf{X}_p\|^q} e_k, \tag{1}$$

where $q = 3$, $\mathbf{N}_p$ is the normalized normal, $\mathbf{L}_k$ is the 3D position of the k-th light, e_k the corresponding intensity, $\mathbf{X}_p$ is the 3D position of a generic point of the surface and finally ρ_p is the albedo, where p denotes the pixel or spatial index. Notice that the intensity fall-off is inversely proportional to the square distance of the light source from the object. In [25] the attenuation term is considered to be inversely proportional to the distance instead of the square distance of the light from the surface point and in this case we have $q = 2$. In this work we investigate both cases ($q = 2$ and $q = 3$).

4 Calibrated Near-Light Photometric Stereo

In this section we propose a solution for the calibrated near-light photometric stereo. By assuming a known focal length f, camera pixel size μ and approximate mean depth $\widehat{z}$ of the object from the camera, we define the relationship between real world coordinates (x, y) and image coordinates (u, v) as

$$(u, v) = \frac{\mu \widehat{z}}{f} \cdot (x, y). \tag{2}$$

Let us write the 3D coordinates and normal of a scene point as

$$\mathbf{X}_p = d\,[x, y, z]^T, \qquad \mathbf{N}_p \propto [\nabla z, -1]^T, \tag{3}$$

where $d \doteq \frac{\mu \widehat{z}}{f}$. Let the light pencil be

$$\widehat{\mathbf{L}}_{pk} \doteq \frac{\mathbf{L}_k - \mathbf{X}_p}{\|\mathbf{L}_k - \mathbf{X}_p\|^q} e_k, \tag{4}$$

where one can notice the dependence from both pixel index p and light index k. Then, we can rewrite the image formation model as

$$I_{pk} = \frac{\mathbf{B}_p^T (\mathbf{L}_k - \mathbf{X}_p)}{\|\mathbf{L}_k - \mathbf{X}_p\|^q} e_k = \mathbf{B}_p^T \widehat{\mathbf{L}}_{pk}, \tag{5}$$

where $\mathbf{B}_p$ are the normals scaled by the albedo. In the calibrated case we know the 3D locations of the point light sources, but we can not determine the *per-pixel* light directions because we have no information about the depth of the scene. For this reason we estimate lights and depth in an iterative way. Initially, we set $\mathbf{X}_p = d\begin{bmatrix} x & y & 1 \end{bmatrix}^T$. Once we have computed the light pencil we can estimate the normals for the distant light case. We do this for each pixel p independently via simple least squares: $\mathbf{B}_p = \left(\widehat{\mathbf{L}}_p^{\dagger} I_p\right)^T$. Then, we integrate the normals [2], update the depth map and repeat. Experimentally we noticed that convergence is fast (about 30-40 iterations). Finally, one can extract the albedo ρ_p from the estimated scaled normals $\mathbf{B}_p$ via $\rho_p = \|\mathbf{B}_p\|_2$, where $\|\cdot\|_2$ denotes the L^2 norm.

5 Uncalibrated Near-Light Photometric Stereo

Let us now consider the uncalibrated case where no information about the 3D locations and intensities of the lights is provided. In the distant light assumption (under orthographic projection) it is well-known that there are ambiguities [7]. One may wonder what ambiguities exist in the near-light case.

Linear ambiguities. This analysis is limited to a linear transformation of the unknowns. From eq. (1), we can transform normals, light positions and points by generic 3×3 matrices Q_N and Q_L and the albedo and light intensities by an arbitrary scaling such that

$$I_{pk} = \frac{\widetilde{\rho}_p \widetilde{\mathbf{N}}_p^T (\widetilde{\mathbf{L}}_k - \widetilde{\mathbf{X}}_p)}{\|\widetilde{\mathbf{L}}_k - \widetilde{\mathbf{X}}_p\|^q} \widetilde{e}_k = \frac{\widetilde{\rho}_p \mathbf{N}_p^T Q_N Q_L (\mathbf{L}_k - \mathbf{X}_p)}{\|Q_L (\mathbf{L}_k - \mathbf{X}_p)\|^q} \widetilde{e}_k. \tag{6}$$

Notice that given a surface point $\widetilde{\mathbf{X}}_p = dQ_L[x,y,z]^T$ and its normal vector $\widetilde{\mathbf{N}}_p = Q_N^T[\nabla z^T, -1]^T$, we obtain that $Q_L = Q_N^{-1}$. Also, one can conclude that the only family of transformations Q_L that does not change $\|\mathbf{L}_k - \mathbf{X}_p\|^q$ are rotations. There can also be a scaling factor s which is absorbed inside the product between $\widetilde{\rho}_p$ and $\widetilde{e}_k$. Indeed, with R a rotation matrix, we have $\widetilde{\mathbf{N}}_p = R^{-T}\mathbf{N}_p = R\mathbf{N}_p, \widetilde{\mathbf{L}}_k = sR\mathbf{L}_k, \widetilde{\mathbf{X}}_p = sR\mathbf{X}_p$ and $\widetilde{\rho}_p\widetilde{e}_k = s^{q-1}\rho_p e_k$. This means that the same observations can be obtained by just rotations and scalings of the lights, the points, and the normals, and by scaling the albedo and light intensities.

Mean depth ambiguity. So far we have stated that the mean depth $\widehat{z}$ is given. Here we show that this parameter needs to be fixed as it constitutes an ambiguity. Firstly, $\mathbf{N}_p$ does not depend on $\widehat{z}$. However, an incorrect $\widehat{z}^* = t\widehat{z}$, for some $t > 0$, will have an effect on the point $\widehat{\mathbf{X}}_p = dt[x,y,z+(t-1)\widehat{z}]^T$. As this error is constant, it can be compensated by the lights by setting $\widehat{\mathbf{L}}_k = t\left(\mathbf{L}_k + d[0,0,(t-1)\widehat{z}]^T\right)$ and by the albedo and light intensities, by setting $\widetilde{\rho}_p\widetilde{e}_k = t^{q-1}\rho_p e_k$. This shows that for any guess of the mean depth the other unknowns are scaled and translated versions of the true parameters.

Alternating minimization algorithm. To obtain a unique solution, one needs to fix at least the above rotation and scaling ambiguities. The rotation is determined by three parameters: the rotation around each of the axis in 3D. We fix two of them by defining the (x,y) reference system of the points to be perpendicular with the viewing direction. By using (x,y) and by assuming that the depth is not a plane we obtain $R = I_d$ (the identity matrix). To fix the remaining parameters it is sufficient to fix d (*i.e.*, the mean depth) and the albedo scale. We calculate the normals, albedo, light positions and light intensities iteratively.

Our alternating minimization procedure is made of two steps where we first estimate depth, albedo and normals by using the calibrated algorithm, and then we estimate the lights and their intensities given the normals, the depth and the albedo. We estimate the light positions and intensities by minimizing the following objective function

$$\min_{\mathbf{L}_k, e_k} F(\mathbf{L}_k, e_k) \doteq \min_{\mathbf{L}_k, e_k} \sum_p \left(I_{pk} - e_k \frac{\mathbf{B}_p^T (\mathbf{L}_k - \mathbf{X}_p)}{\|\mathbf{L}_k - \mathbf{X}_p\|^q} \right)^2. \tag{7}$$

Since the cost functional is non convex, we perform an initial brute force coarse search in order to get an initial estimate for the light positions. To make the search computationally

initialization
$\mathbf{X}_p \leftarrow d\left[x, y, 1\right]^T, \mathbf{B}_p \leftarrow \left[0, 0, -1\right]^T$;
$(\mathbf{L}_k, e_k) \leftarrow \min_{\mathbf{L}_k, e_k} \sum_p \left(I_{pk} - e_k \frac{\mathbf{B}_p^T (\mathbf{L}_k - \mathbf{X}_p)}{\|\mathbf{L}_k - \mathbf{X}_p\|^q} \right)^2$;
repeat
  $\widehat{\mathbf{L}}_{pk} \leftarrow e_k (\mathbf{L}_k - \mathbf{X}_p) / \|\mathbf{L}_k - \mathbf{X}_p\|^q$;
  $\mathbf{B}_p \leftarrow \left(\widehat{\mathbf{L}}_p^{\dagger} I_p \right)^T$;
  $z \leftarrow$ Poisson Integration$(\mathbf{B}_p)$ [2] ;
  $\mathbf{X}_p \leftarrow d \begin{bmatrix} p \\ z \end{bmatrix}$;
  $(\mathbf{L}_k, e_k) \leftarrow \min_{\mathbf{L}_k, e_k} \sum_p \left(I_{pk} - e_k \frac{\mathbf{B}_p^T (\mathbf{L}_k - \mathbf{X}_p)}{\|\mathbf{L}_k - \mathbf{X}_p\|^q} \right)^2$;
until $\text{mean}(z^{i+1} - z^i) < 10^{-4} cm$;

Algorithm 1: Uncalibrated Near-Light PS Algorithm

efficient we search within a one meter cubic volume with a 10 cm step for the light positions. For every light candidate we obtain a closed form solution for the light intensities

$$e_k = \sum_p \left(I_{pk} \frac{\mathbf{B}_p^T (\mathbf{L}_p - \mathbf{X}_p)}{\|\mathbf{L}_k - \mathbf{X}_p\|^q} \right) / \sum_p \left(\frac{\mathbf{B}_p^T (\mathbf{L}_k - \mathbf{X}_p)}{\|\mathbf{L}_k - \mathbf{X}_p\|^q} \right)^2. \tag{8}$$

Second, we refine for the light positions via gradient descent $\mathbf{L}_k^{i+1} = \mathbf{L}_k^i - \gamma \frac{\partial F}{\partial \mathbf{L}_k}$, where

$$\frac{\partial F}{\partial \mathbf{L}_k} = -2 \sum_p \left(I_{pk} - e_k \frac{\mathbf{B}_p^T (\mathbf{L}_k - \mathbf{X}_p)}{\|\mathbf{L}_k - \mathbf{X}_p\|^q} \right) \cdot \left(e_k \frac{\mathbf{B}_p \|\mathbf{L}_k - \mathbf{X}_p\|^2 - q(\mathbf{L}_k - \mathbf{X}_p)^T \mathbf{B}_p (\mathbf{L}_k - \mathbf{X}_p)}{\|\mathbf{L}_k - \mathbf{X}_p\|^{q+2}} \right) \tag{9}$$

and $\gamma > 0$ is a small step coefficient. The steps are summarized in Algorithm 1.

6 Experiments

We performed real world experiments to validate our method and show the advantages of the near-light model compared to the distant light one. We acquired images of four real world scenes: **Dwarf**, **Sphere**, **Face** and **House**, which we will make publicly available. For the first two datasets (**Dwarf**, **Sphere**) we captured images with controlled illumination where the camera was positioned 60 cm above the scene. We randomly distributed 12 led lights in the upper hemisphere of the scene, which were positioned within a distance range of 40-60 cm. The light calibration was done manually (in order to have a ground truth reference) and the error is less than 0.5 cm. We also used the camera calibration toolbox from Bouguet [3] in order to calculate the rotation and translation that bring the light coordinates onto the camera reference system. For the second two datasets (**Face** and **House**) we used a second experimental setup where lights where not calibrated. The images were captured (20 for the **Face** and 15 for the **House** dataset) by simply moving arbitrarily a light bulb around and close to the surface of the object or face. We only made sure that the lights were placed between the camera and the scene. The minimum number of light sources is 3. However,

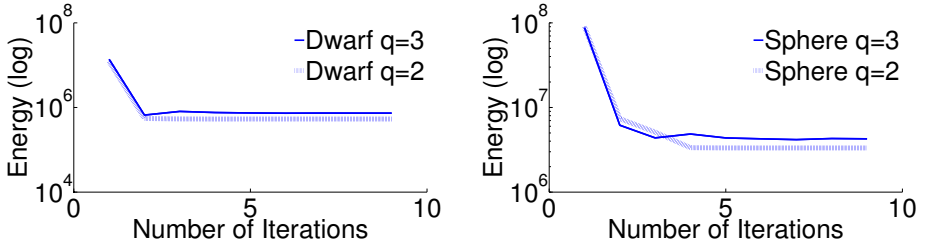

Figure 1: Fitting error (y-axis) of the model to the observations for the **Dwarf** (left) and **Sphere** (right) dataset for two different choices of the coefficient q (see eq. (1)). On the x-axis we show the number of iterations required for the energy cost function to converge. However, the convergence of the estimated normals and lights requires more than 8 iterations (typically 30-40 iterations).

we used more of them in order to make the method more robust against noise, shadows, specularities and interreflections. For both setups we used a Canon 5D Mark III with a 50 mm lens fixed on a tripod.

In Fig. 3 we show the experimental results in the case of the **Dwarf** and **Sphere** datasets. We have included additional profile photos in order to create a better perception of the 3D structure of the scene. In Fig. 1 we plot the data fitting term

$$\text{Energy} \doteq \sum_{p,k} \left(I_{pk} - e_k \frac{\mathbf{B}_p^T (\mathbf{L}_k - \mathbf{X}_p)}{\|\mathbf{L}_k - \mathbf{X}_p\|^q} \right)^2 \tag{10}$$

for $q = 2,3$ against the iteration number. It can be observed that the algorithm converges quite quickly, although a small variation in the data fitting can result in a large variation of the depth estimate. Indeed, the depth estimate takes about 30-40 iterations to converge (changes between two consecutive estimates is within 0.1% of one of them). Moreover, it appears that for these datasets, a choice of $q = 2$ in the image formation model yields a lower fitting term, compared to that for $q = 3$. Indeed, the reconstruction results in the first case are better (see Fig. 3). The light estimation is more accurate as well: for the **Dwarf** dataset we obtain a mean error in the light coordinates estimation of 4.79 cm for $q = 2$ and 5.35 cm for $q = 3$, while for the **Sphere** dataset such error is 3.85 cm for $q = 2$ and 5.25 cm for $q = 3$. We also performed the reconstruction of a ground truth scene (planar scene made of paper) and obtained a mean angular error in the surface normals estimation of 4.05 angular degrees for $q = 2$ and 9.47 angular degrees for $q = 3$, while the distant light photometric stereo method gives a mean angular error of 24.85 angular degrees. These results seem to be in contradiction with the well established image formation model for near-light illumination which requires $q = 3$. This might be due to the light sources we chose for the illumination setup. However, for both cases the reconstruction results obtained with our method are very good. Indeed, notice the significant improvement of the reconstruction compared to the distant light photometric stereo. Conventional photometric stereo fails because the lights are close to the scene and the distant light assumption does not hold anymore and a strong distortion of the normal map can be noticed, especially towards the borders of the image. Our uncalibrated method performs better and the solution is very close to the calibrated near-light case. However, the surface is smoothed out at the borders of the object. This is because of the shadows which introduce non-negligible distortion to the imaging model,

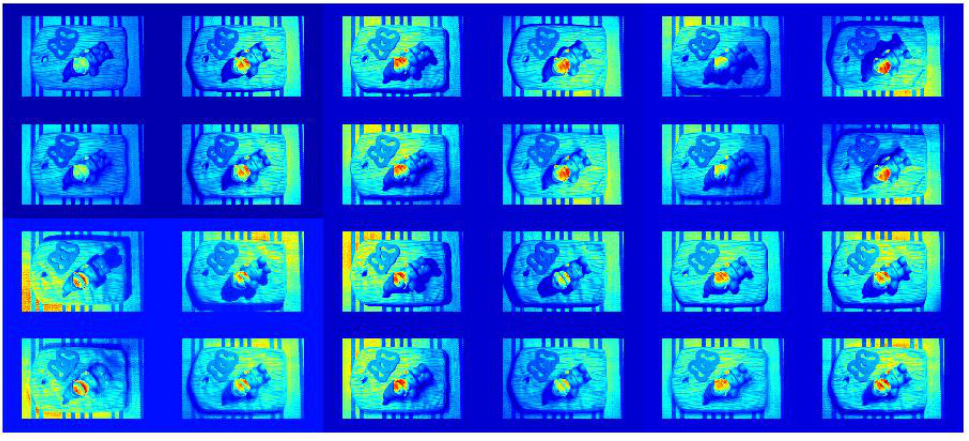

Figure 2: Fitting of our model to real observations in the **Dwarf** dataset (best viewed in color). There are twelve pairs of images where for each pair, on the top (first and third row from the top) we show an input image, and below (second and fourth row from the top) the corresponding synthetic one generated with our model and the parameters estimated via our uncalibrated near-light PS algorithm. The synthetic images are almost identical to the captured ones except at the shadow areas.

especially when the lights are closer to the scene, as in our experimental setup. Moreover, the effect of interreflections at these regions with strong concave edges is significant. In Fig. 2 we also compare the synthetic images generated with our model (for $q = 2$) and estimated parameters with the input images of the **Dwarf** dataset. Notice that most of the intensities are well modeled except for shadows.

In Fig. 4 we show the experimental results in the case of the **Face** and **House** datasets for a choice of $q = 2$. In this case we included two additional photos for the ground truth perception. We compare our uncalibrated distant light PS method with the state-of-art in uncalibrated distant light PS, the Total Variation method [23] and the Diffuse Maxima one [20]. Again, one can notice really accurate details in the reconstruction and the improvement with respect to the distant lights case. Finally, the running time of our algorithm for the above datasets with resolution 0.2-0.3 megapixels varies between 3 and 4 minutes.

7 Discussion

In near-light photometric stereo (both calibrated and uncalibrated case) shadows are a problem and need to be handled properly. Indeed, the reconstruction errors due to the shadows can be noticed under the nose of the **Face** or at the edges of the **Dwarf** toy and result in over smoothing of the depth map (see Fig. 5). However, dealing with shadows in the uncalibrated near-light scenario, where subspace clustering techniques such as [27] are not applicable, is not a trivial task. It would be interesting to model also interreflections at these shadow regions. This would drastically improve the reconstruction results for many indoor scenes or underground navigation scenarios where shadows may be dominant. This study will be the focus of our future work.

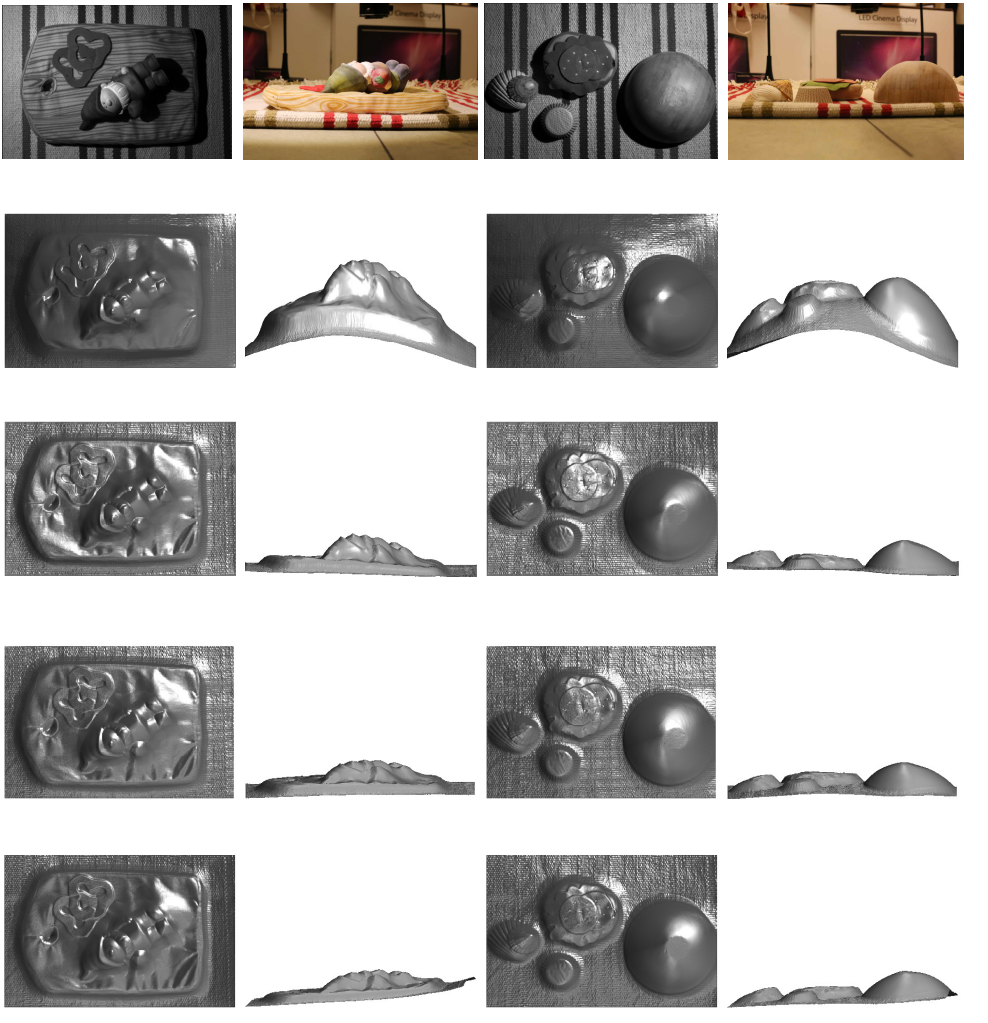

Figure 3: Reconstruction results for the **Dwarf** and **Sphere** scene obtained via the second experimental setup. Rows from top to bottom: frontal (first and third column from left) and lateral (second and fourth column from left) view of the scene, reconstructed surfaces via calibrated distant light PS (second row from top), reconstructed surfaces via our calibrated near-light PS (third row from top), reconstructed surface via our uncalibrated near-light PS method with $q = 2$ (fourth row from top) and reconstructed surface via our uncalibrated near-light PS method with $q = 3$ (fifth row from top).

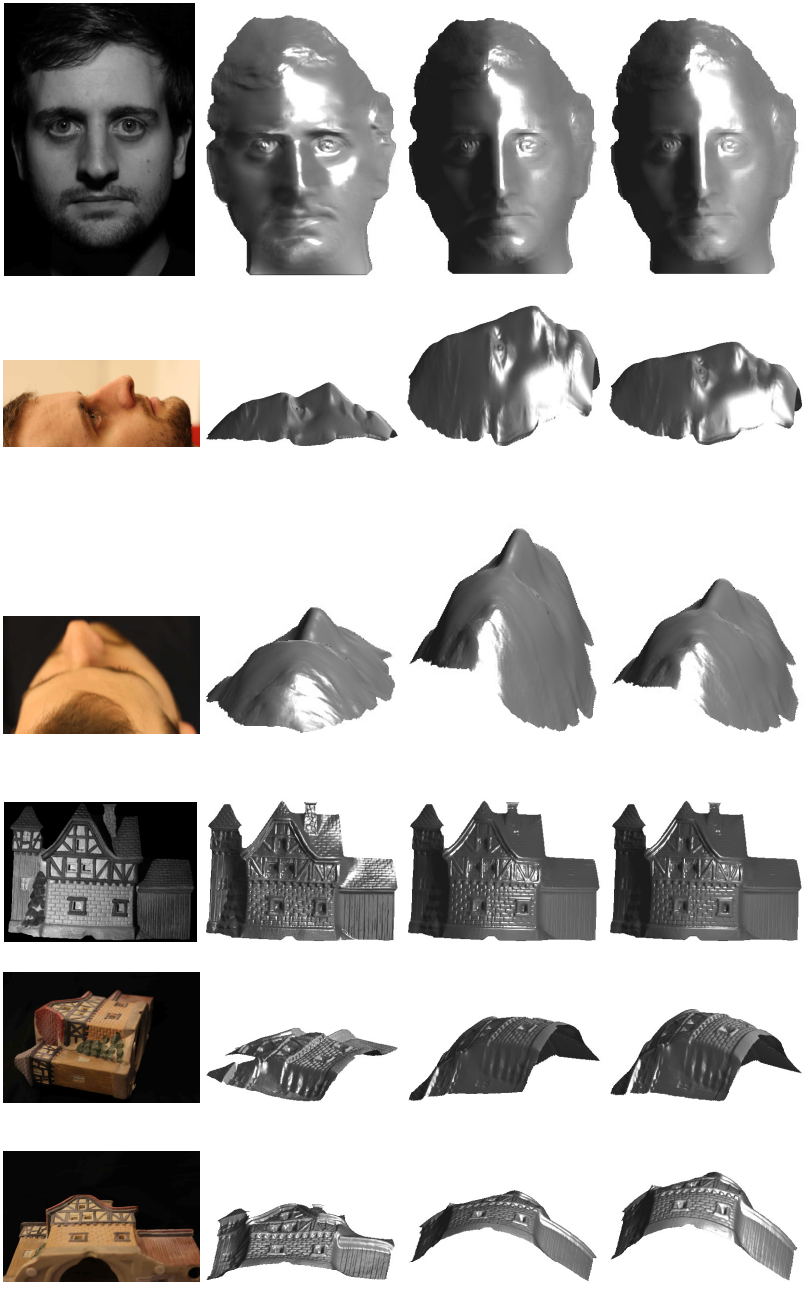

Figure 4: Reconstruction results for the **Face** and **House** dataset obtained via the second experimental setup. Columns from left to right: one input image (first and fourth row from top) and two views of each scene (second, third, fifth and sixth row from top), reconstructed surface via our uncalibrated near-light PS method (second column from left), the estimated surfaces via [20] (third column from left) and the estimated surfaces via [23] (fourth column from left).

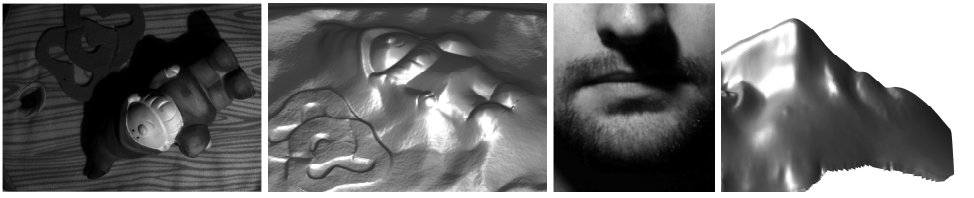

Figure 5: Surface reconstruction errors due to shadows and interreflections. They can be noticed around the borders for the **Dwarf** and under the nose for the **Face** rendered surface.

8 Conclusion

In this work we investigated the uncalibrated near-light photometric stereo problem. Our method can handle arbitrary spatially variant unknown illumination and does not make any assumptions about the geometry, lights or reflectance of the scene. We proposed an iterative algorithm which converges quickly to the true solution. We validated our method on real world experiments and showed that significant improvement in the surface reconstruction can be achieved compared to the distant-light case.

References

[1] Jens Ackermann, Simon Fuhrmann, and Michael Goesele. Geometric Point Light Source Calibration. In *Proceedings of Vision, Modeling and Visualization*, 2013.

[2] Amit Agrawal and Ramesh Raskar. What is the range of surface reconstructions from a gradient field. In *In ECCV*, pages 578–591. Springer, 2006.

[3] Jahanzeb Ahmad, Jiuai Sun, Lyndon Smith, and Melvyn Smith. An improved photometric stereo through distance estimation and light vector optimization from diffused maxima region. *Pattern Recognition Letters*, 2013.

[4] N.G. Alldrin, S.P. Mallick, and D.J. Kriegman. Resolving the generalized bas-relief ambiguity by entropy minimization. *Computer Vision and Pattern Recognition*, pages 1 –7, 17-22 2007.

[5] Robert Anderson, Björn Stenger, and Roberto Cipolla. Augmenting depth camera output using photometric stereo. In *MVA*, pages 369–372, 2011.

[6] T. Aoto, T. Taketomi, T. Sato, Y. Mukaigawa, and N. Yokoya. Position estimation of near point light sources using a clear hollow sphere. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 3721–3724, 2012.

[7] Peter N. Belhumeur, David J. Kriegman, and Alan L. Yuille. The bas-relief ambiguity. *Int'l Journal of Computer Vision*, pages 35(1):33–44, 1999.

[8] J. Y. Bouguet. Camera calibration toolbox for matlab. In *Technical report: http://www.vision.caltech.edu/bouguetj/calib doc/*, 2007.

[9] Manmohan Krishna Chandraker, Fredrik Kahl, and David J. Kriegman. Reflections on the generalized bas-relief ambiguity. In *Conference on Computer Vision and Pattern Recognition*, pages 788–795, 2005.

[10] J.J. Clark. Active photometric stereo. In *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR '92., 1992 IEEE Computer Society Conference on*, pages 29–34, 1992.

[11] Toby Collins and Adrien Bartoli. 3d reconstruction in laparoscopy with close-range photometric stereo. In *MICCAI (2)*, pages 634–642, 2012.

[12] Microsoft Corp.RedmondWA.KinectforXbox360.

[13] O. Drbohlav and M.J. Chantler. Can two specular pixels calibrate photometric stereo? *Proc. of Int'l Conf. on Computer Vision*, pages 1850–1857, 2005.

[14] K. Hara, K. Nishino, and K. lkeuchi. Light source position and reflectance estimation from a single view without the distant illumination assumption. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(4):493–505, 2005. ISSN 0162-8828.

[15] H. Hayakawa. Photometric stereo under a light-source with arbitrary motion. *Journal of the Optical Society of America*, 11(11):3079–3089, 1994.

[16] Tomoaki Higo, Yasuyuki Matsushita, Neel Joshi, and Katsushi Ikeuchi. K.: A hand-held photometric stereo camera for 3-d modeling. In *In: Proc. ICCV. (2009)*.

[17] Byungil Kim and Peter Burger. Depth and shape from shading using the photometric stereo method. *CVGIP: Image Understanding*, 54(3):416 – 427, 1991.

[18] Sanjeev Jagannatha Koppal and Srinivasa G. Narasimhan. Novel depth cues from uncalibrated near-field lighting. *ICCV, 2007*.

[19] Ryszard Kozera and Lyle Noakes. Noise reduction in photometric stereo with non-distant light sources. In K. Wojciechowski, B. Smolka, H. Palus, R.S. Kozera, W. Skarbek, and L. Noakes, editors, *Computer Vision and Graphics*, volume 32 of *Computational Imaging and Vision*, pages 103–110. 2006.

[20] Thoma Papadhimitri and Paolo Favaro. A closed-form, consistent and robust solution to uncalibrated photometric stereo via local diffuse reflectance maxima. *International Journal of Computer Vision*, pages 1–16, 2013. ISSN 0920-5691.

[21] Thoma Papadhimitri and Paolo Favaro. A new perspective on uncalibrated photometric stereo. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1474–1481, 2013.

[22] Mark W. Powell, Student Member, Sudeep Sarkar, Dmitry Goldgof, and Senior Member. A simple strategy for calibrating the geometry of light sources. *PAMI*, 23:1022–1027, 2001.

[23] Yvain Quéau, François Lauze, and Jean-Denis Durou. Solving the uncalibrated photometric stereo problem using total variation. In *SSVM*, pages 270–281, 2013.

[24] HU Rashid and P Burger. Differential algorithm for the determination of shape from shading using a point light source. *Image and Vision Computing*, 10(2):119 – 127, 1992. ISSN 0262-8856.

[25] Fumihiko Sakaue and Jun Sato. A new approach of photometric stereo from linear image representation under close lighting. In *ICCV Workshops*, pages 759–766, 2011.

[26] B. Shi, Y. Matsushita, Y. Wei, Ch. Xu, and P. Tan. Self-calibrating photometric stereo. *In IEEE Conference on Computer Vision and Pattern Recognition*, pages 1118–1125, 2010.

[27] Kalyan Sunkavalli, Todd Zickler, and Hanspeter Pfister. Visibility subspaces: uncalibrated photometric stereo with shadows. In *Proceedings of the 11th European conference on Computer vision: Part II*, ECCV'10, pages 251–264, 2010.

[28] T. Takai, K. Niinuma, A. Maki, and T. Matsuyama. Difference sphere: an approach to near light source estimation. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 1, pages I–98–I–105 Vol.1, 2004.

[29] Martin Weber and Roberto Cipolla. A practical method for estimation of point light-sources. *PROCEEDINGS BMVC 2001*, 2:471–480, 2001.

[30] Kwan-Yee K. Wong, Dirk Schnieders, and Shuda Li. Recovering light directions and camera poses from a single sphere. ECCV '08, pages 631–642, 2008. ISBN 978-3-540-88681-5.

[31] R.J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.

[32] Chenyu Wu, Srinivasa G Narasimhan, and Branislav Jaramaz. A multi-image shape-from-shading framework for near-lighting perspective endoscopes. *International Journal of Computer Vision*, February 2009.

[33] Lei Yang, Shiwei Ma, and Bo Tian. New shape-from-shading method with near-scene point lighting source condition. In Yinglin Wang and Tianrui Li, editors, *Foundations of Intelligent Systems*, volume 122 of *Advances in Intelligent and Soft Computing*, pages 653–664. 2012.

[34] Qing Zhang, Mao Ye, Ruigang Yang, Yasuyuki Matsushita, Bennett Wilburn, and Huimin Yu. Edge-preserving photometric stereo via depth fusion. In *CVPR*, pages 2472–2479. IEEE, 2012.