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

Photometric stereo (PS) [3] is a technique to accurately recover the normal map of a 3D scene from several pictures (at least three) taken from the same view point 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. Motivated by this fact, we introduce for the fist time an uncalibrated near-light photometric stereo method where no prior information about light position and intensities is needed. Only in [1] 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 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.

The image formation model typically used for the near-light case under the Lambertian reflectance is

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

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 [2] 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 $q = 3$).

We solve the uncalibrated near-light photometric stereo via an alternating minimization procedure which consists of two steps: first we estimate the normals, the albedo and the depth and then we estimate the lights and their intensities given the normals, the depth and the albedo.

In Fig. 1 we show the experimental results in the case of the **Dwarf** and **Sphere** datasets. We captured images by randomly distributing 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 have included additional profile photos in order to create a better perception of the 3D structure of the scene. It can be noticed that a choice of $q = 2$ in the image formation model yields to lower reconstruction errors compared to that for $q = 3$. 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.

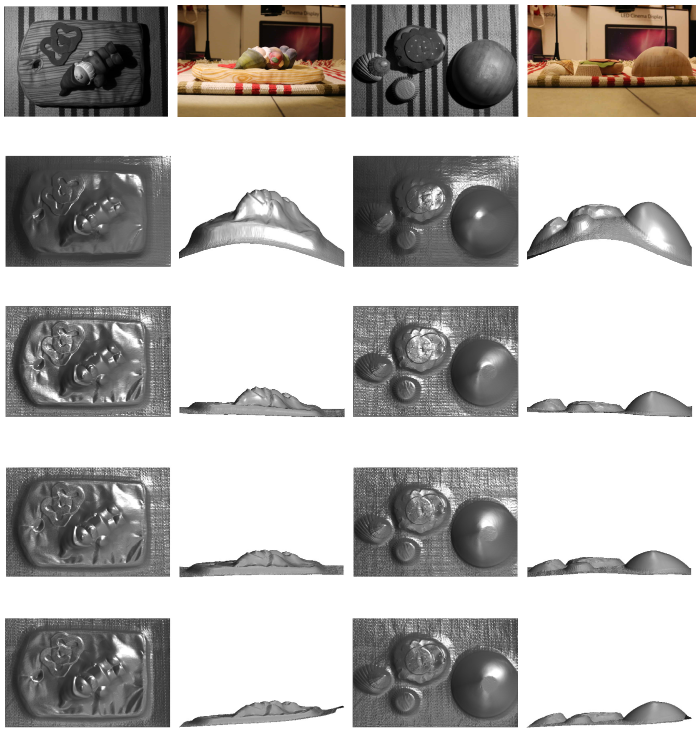

Figure 1: Reconstruction results for the **Dwarf** and **Sphere** scene obtained via our experimental setup. Rows from top to bottom: frontal (first and third column from left) and lateral (second and fourth column from left) view of the scene, reconstructed surfaces via calibrated distant light PS (second 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).

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, 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. Finally, the running time of our algorithm for the above datasets with resolution 0.2-0.3 megapixels varies between 3 and 4 minutes.

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

[2] 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.

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