LUCES: A Dataset for Near-Field Point Light Source Photometric Stereo

Roberto Mecca[1]
rmecca@crl.toshiba.co.uk
Fotios Logothetis[1]
flogothetis@crl.toshiba.co.uk
Ignas Budvytis[2]
ib255@cam.ac.uk
Roberto Cipolla[2]
rc10001@cam.ac.uk

[1] Cambridge Research Laboratory, Toshiba Europe, Cambridge, UK Cambridge, UK
[2] University of Cambridge Cambridge, UK

Abstract

Three-dimensional reconstruction of objects from shading information is a challenging task in computer vision. As most of the approaches facing the Photometric Stereo problem use simplified far-field assumptions, real-world scenarios have essentially more complex physical effects that need to be handled for accurately reconstructing the 3D shape. An increasing number of methods have been proposed to address the problem when point light sources are assumed to be nearby the target object.

To understand the capability of the approaches dealing with this near-field scenario, the literature till now has used synthetically rendered photometric images or minimal and very customised real-world data. In order to fill the gap in evaluating near-field photometric stereo methods, we introduce LUCES the first real-world 'dataset for near-fieLd point light soUrCe photomEtric Stereo' of 14 objects of different materials. 52 LEDs have been used to lit each object positioned 10 to 30 centimeters away from the camera. Together with the raw images, in order to evaluate the 3D reconstructions, the dataset includes both normal and depth maps for comparing different features of the retrieved 3D geometry. Furthermore, we evaluate the performance of the latest near-field Photometric Stereo algorithms on the proposed dataset to assess the state-of-the-art method with respect to actual close range effects and object materials.

1 Introduction

Since the introduction of the Photometric Stereo problem (PS) by Woodham in the early '80s [40], a wide variety of approaches tackled the very same problem of reconstructing 3D geometry of an object under varying illumination from the same view point. Despite the very simplified assumption in [40] to make the PS problem solvable as an (over-determined) linear system, similar simplifications are often still considered nowadays to make the problem applicable to real-world scenarios. Nonetheless, diffuse material assumption was relaxed in [15, 36, 39], camera perspective viewing was modelled in [23, 38], and robust optimisation

Figure 1: From left to right:(1) the stage of our Photometric Stereo setup (2) a top view of a sample object (Squirrel), (3) acquisition with the GOM scanner (4) the 3D scanned mesh.

methods were employed by [10, 14] to increase robustness to outliers. Light calibration assumption was also relaxed by [27, 29].

One of the most challenging aims of more recent PS methods is realistic illumination modelling, as uniform directional lighting is hard to achieve in practice. For this purpose, several methods have proposed using point light sources instead of directional ones [8, 16, 18, 20, 24, 26, 32]. As LED illumination based technology has spread widely, point light source has become by far the more adopted alternative to directional lighting. However, point light sources require non-linear modeling of light propagation and attenuation, but they are a more realistic assumption than directional lights for near-field photometric imaging acquisitions. Note that proximity of the camera and lights to the object are very favorable in order to capture detailed geometry and minimise the ambient light interference. For example, near-field photometric stereo has been used in practice with handheld acquisition devices [12] and in endoscope-like inspections [9]. Whereas, the far-field assumptions do not allow to combine PS with multi-view for volumetric reconstruction [19].

However, despite the increased contribution from the computer vision community to tackle the near-field PS problem, the evaluation of such methods has relied on synthetic [20] or very minimal real-world datasets [32, 35]. The lack of shared data has prevented detailed and fair comparisons across the different methods. The aim of this work is to provide a comprehensive near-field PS benchmark with a variety of objects having different materials in order to evaluate several algorithms and understand their strengths and weaknesses. For this purpose, ground truth normal map and depth are provided for each object.

Our contribution is as follows:

- introducing the first near-field PS dataset of 14 real objects having a wide variety of materials;
- evaluating most relevant algorithms for the near-field PS problem and establish the actual state-of-the-art method.

The dataset (including all images, light and camera calibration parameters and ground truth meshes) and the evaluation of the methods are available for download at: http://www.robertomecca.com/psdataset.html.

2 Related Work

A number of approaches for the PS problem has been proposed since it was first introduced [40]. We refer to some fairly recent surveys [3, 11] to cover the initial evolution of the PS

Figure 2: Top view of the objects captured for this dataset. Below every object the acquisition distance between the object and the camera, and the material of the object are reported.

methodologies. Here we discuss more contemporary algorithms as this work focuses on evaluating their performances on the proposed dataset.

2.1 PS datasets

Across the years, a number of custom real-world PS datasets have been created to suit the purposes of the proposed approaches. Alldrin *et al.* [4] proposed a dataset consisting of 3 objects lit by roughly a hundred distant light directions. The light calibration in terms of positioning and intensity has been performed by using respectively a mirror sphere and a diffuse sphere. Xiong *et al.* [41] have proposed a dataset of 7 objects using 20 directional lights calibrated with two chrome spheres. As the approach was mostly modeling PS images with Lamberitan irradiance equations, the material of the objects was quite diffuse. A limited number of PS data has been released by Quéau *et al.* to prove the working principle of an edge preserving method [28] and a multi-spectral PS approach [30].

Although initially designed for evaluating multi-view approaches, the datasets released by Aanæs *et al.* [1, 2] are useful for evaluating PS approaches as they also contain images under varying illumination.

As most of the methods aimed at tackling the PS problem deal with the far-field setting, recently Shi *et al.* [36] introduced the first dataset in this category, namely DiLiGenT aimed at evaluating reconstruction methods over a wide variety of materials for 10 different objects. This work also contains a well discussed taxonomy for non-Lambertian and uncalibrated PS approaches. Their setup consists of 96 LEDs placed several meters away from the objects to approximate directional illumination and the camera (with a 50mm lens) was placed at 1.5m from the object. Such distance between the object and the camera/lights system does not provide to this dataset the near-field light variation studied in many recent approaches.

Near-field datasets: There are very limited, proper near-field labeled data including a single object from [32] and 3 simple objects from [35].

2.2 Near-field PS

The near-field setting is intrinsically more complicated to model than the far-field one as it requires handling not only different type of BRDFs [5, 25] but also anisotropic light propagation [26], inconsistent light intensity among the set of LEDs [18, 31] and finally the uncalibrated case [27]. Given this wide variety of difficulties, most of the proposed near-field PS methods have presented custom PS data.

With the aim to tackle all these issues simultaneously for the near-field PS, latest approaches have been exploiting deep learning capability training their networks with synthetically rendered data and data driven rendered data [22]. In particular, Logothetis *et al.* [20] used a per-pixel training strategy that allows to render unlimited data without carrying any training dataset. The sim-to-real gap is then filled by augmenting the data with physical effects such as noise, ambient light, interreflections, etc. Finally, the far to near field compensation is performed by integrating the normal field to compute the depth. By doing so iterativelly, the method converges to an estimate of the 3D geometry. Santo *et al.* [35] have recently introduced a near-field PS method where the near-field compensation is computed after computing the far-field normals map from PSFCN [6]. The surface optimisation is performed through a differentiable renderer which fuses the normal predictions and the lighting model to re-project to the original images. This step limits the evaluation of the method to small images due to very high requirements of GPU RAM (around 20GB for 0.5Mpx images). Furthermore, despite the near-field setting, the camera viewing is assumed orthographic.

3 Data Capture

This section gives an overview of the data capture and calibration procedure.

3.1 Photometric Stereo Data Capture

The Photometric Stereo setup. Our setup (see Figure 1, left) consists of the following main components:

- RGB camera FLIR BFS-U3-32S4C-C with 8mm lens
- 52 LED Golden Dragon OSRAM
- variable voltage for adjustable LED power
- Arduino Mega 2560

A custom printed circuit board (PCB) has been designed to host 52 bright LED controlled with by an Arduino Mega. The configuration of the LEDs was planar around the camera. A set of 52 images was captured per object. The camera parameters (aperture and shutter speed) and LED voltage were adjusted to achieve the best object exposure, which is very critical for specular objects. In particular, ISO sensitivity was set to zero and the exposure time has been changed depending on the shininess of the object (between 9 and 500 ms). We also changed the power of the LEDs for particularly specular objects to avoid over saturated images. We used the maximum color-depth possible for the camera which was 12-bit. All camera prepossessing was turned off during the acquisition, including white-balance and analog gain.

Figure 3: Demonstration of the processing steps performed per object. Firstly, compensation for radial distortion and demosaicing is performed on raw images to get RBG ones (left). Laser-scanned ground truth meshes are aligned with RGB images and ground truth normal maps are rendered (middle). Segmentation masks are generated (removing the pixels corresponding to markers)(right).

Several optomechanical tools have been used for holding the camera and the PCB jointly. A manual XYZ translation stage with differential adjusters has been used to positioning the camera accurately through the printed circuit board.

In order to limit interreflections and ambient light, the walls surrounding the setup have been covered with black, polyurethane-coated nylon fabric.

Camera Intrinsics. This is performed using 100 checkerboard images and the OpenCV calibration toolbox. Fourth degree radial distortion is estimated and this is used to rectify all the images. The calibration re-projection error was 0.42px. The RAW data (before demosaicing and rectification) will also be made available.

Near Lighting Model. The lighting model is the anisotropic point light sources [24] which is used for all SOTA methods ([18, 20, 32, 35] evaluated in Section 4). This model assumes that a light source m, has a position $\mathbf{P}_m \in \mathbb{R}^3$, principal direction $\mathbf{D}_m \in \mathbb{R}^3$, RGB brightness $\Phi_m \in \mathbb{R}^3$ and angular dissipation factor $\mu_m \in \mathbb{R}$. Therefore, a point $\mathbf{X} \in \mathbb{R}^3$ has a lighting vector $\mathbf{L}_m(\mathbf{X}) = \mathbf{P}_m - \mathbf{X}$ and assuming $\hat{\mathbf{L}}_m = \frac{\mathbf{L}_m}{||\mathbf{L}_m||}$ as the normalised light direction, we consider the following light attenuation:

$$a_m(\mathbf{X}) = \frac{(\hat{\mathbf{L}}_m(\mathbf{X}) \cdot \hat{\mathbf{D}}_m)^{\mu_m}}{||\mathbf{L}_m(\mathbf{X})||^2}. \tag{1}$$

Light Calibration. The aim here is not only to estimate the point light position $\mathbf{P}_m$ [34], but also the other LED parameters $\mathbf{D}_m$, Φ_m and μ_m. Instead of employing methods that aim at estimating these parameters while reconstructing the geometry [18, 31, 32], we developed a custom method that accurately estimates $\mathbf{P}_m$, $\mathbf{D}_m$, Φ_m and μ_m from PS images of a purely diffuse reflectance plane. To do so, we used a plane with 99% nominal reflectance in UV-VIS-NIR wavelength range (350 - 1600nm). To have an initial estimate of Φ_m, we measured the LED brightness with a LuxMeter.

For every object, the calibration plane was captured twice, at different distances, in order to get data redundancy and produce a more accurate calibration. Thus, the Lambertian calibration object with albedo ρ and surface normal $\mathbf{N}$, should satisfy the resulting image irradiance equation:

$$\mathbf{I}_m = \Phi_m a_m \rho \hat{\mathbf{L}}_m \cdot \hat{\mathbf{N}}. \tag{2}$$

The irradiance Equation 2 was implemented into a differentiable renderer (using Keras of Tensorflow v2.0) with the LED parameters being the model weights thus allowing refinement from a reasonable initial estimate. The parameters were initialised as follows: Φ_m from the LuxMeter, $\mathbf{D}_m = [0,0,1]$, $\mu_m = 0.5$, $\mathbf{P}_m$ from the schematic of the printed circuit board of the LEDs and $\rho = 1$. We used L_1 loss function for 30 epochs and converged to around 0.005 error i.e 0.5% of the maximum image intensity. The complete calibration parameters are included in the dataset.

3.2 3D Ground Truth Capture

3D capturing device. 3D ground-truth has been acquired with the optical 3D scanner GOM ATOS Core 80/135 with a reported accuracy of 0.03mm (see Figure 1). The GOM scanner uses a stereo camera set-up and more than a dozen scans were performed and fused per object. In order to keep the geometry of the object consistent with the PS data, no spray coating has been used to ease the acquisition. Indeed, coating material can fill up those regions of the objects that are prone to interreflection and that are noticeably harder to reconstruct. Instead, markers were used for some objects.

Alignment. The laser scans of the objects were aligned and merged using MeshLab [7]. Some manual removal of noisy regions was performed and finally screened Poisson reconstruction [17] was used in order to obtain full continuous surfaces (which are both useful for rendering normal maps and for mutual information alignment). As expected, not all parts of the surfaces of all objects have the same amount of noise, especially the metallic objects (Bell, Cup). Meshes were aligned with the photometric stereo images following the same procedure as in [37]. This involved manual initialisation and then refinement using the mutual information registration filter of MeshLab. This was performed repeatedly until the projection of the mesh on the images was visually 'pixel perfect' (using the semi-transparent overlay). Using the aligned meshes, ground truth normal maps were rendered (using Blender). In addition, manual segmentation was performed to remove regions where the GT was unreliable (markers on the objects, holes etc), those masks are provided in the dataset. Furthermore, the dataset contains meshes that have been interpolated in the marker/hole regions. The steps per object are summarised in Figure 3.

3.3 Dataset Overview

For each of the 14 object, 52 PS images have been acquired using the BayerRG16 RAW format. The total amount of PS images amounts then to 728. For all objects, rectified RGB PS images will be released (by compensating for the radial distortion). We note that color balancing was not performed on the images as this will distort the saturated pixels (which is an important feature for CNN-based PS methods [13, 21]). Instead, RGB light source brightness are provided along with the rest of point light source parameters. Both normal map and depth ground truth will be provided in order to evaluate the accuracy of near-field PS methods with either cases.

4 Experiments

In this section, we evaluate four competing near-field methods namely [18, 20, 32, 35]. In addition, we also evaluate with [13], the best performing far-field method (on the far-field benchmark [36]) to demonstrate the need for a near-field method.

Method	Error	Bell	Ball	Buddha	Bunny	Die	Hippo	House	Cup	Owl	Jar	Queen	Squirrel	Bowl	Tool	Average
L17-[18]	MAE	28.25	**9.77**	11.5	20.15	11.95	15.42	29.69	30.76	13.77	10.56	13.05	15.93	12.5	**15.1**	17.03
	MZE	4.45	0.81	4.67	7.51	4.58	3.19	6.99	2.67	3.64	6.56	**1.89**	1.82	4.37	**3.25**	4.02
Q18-[32]	MAE	25.8	12.12	14.07	13.73	13.77	18.51	30.63	37.63	14.74	15.66	13.16	**14.06**	11.19	16.12	17.94
	MZE	12.03	2.5	9.28	7.06	5.91	6.8	8.02	4.83	5.83	16.87	6.92	2.55	6.48	6.69	7.27
S20-[35]	MAE	**9.5**	25.42	19.17	12.5	**5.23**	23.12	**28.02**	**14.22**	13.08	9.27	16.62	14.07	12.44	17.42	15.72
	MZE	1.9	5.5	5.53	6.02	**2.76**	7.04	**6.15**	**1.62**	3.75	6.09	3.91	2.81	5.22	4.68	4.5
L20-[20]	MAE	14.74	12.43	**10.73**	**8.15**	6.55	**7.75**	30.03	23.35	**12.39**	**8.6**	**10.96**	15.12	**8.78**	17.05	**13.33**
	MZE	**1.53**	**0.67**	**3.27**	**2.49**	4.44	**1.82**	9.14	2.04	**3.44**	**3.86**	1.94	**1.01**	**2.80**	5.90	**3.17**
I18-[13]	MAE	23.55	44.29	35.29	36	41.52	44.9	49.05	35.78	40.27	40.66	32.89	41.09	28.04	31.71	37.5
	MZE	5.93	6.59	10.92	6.88	7.83	7.59	8.98	3.17	8.67	15.54	8.08	5.8	6.69	12.45	8.22
GT	Diff-MAE	2.5	2.69	2.69	2.93	2.49	3.2	9.19	2.85	4.3	1.79	4.22	3.26	2.27	2.34	3.34
	Int[28]-MZE	0.08	0.22	3.28	2.30	0.56	1.28	7.43	0.02	3.51	0.12	3.25	1.12	0.12	0.13	1.67

Table 1: Complete evaluation of five methods on all objects. Mean angular error MAE (degrees) and mean depth error MZE (mm) are reported. The last two lines contain the error obtained after differentiation and integration of GT depth and normals respectively.

Evaluation hyper-parameters. [18, 32] and [35] have publicly available code whereas for [20] the code has been provided. Indeed, [20] has the disadvantage that the light configurations has to be known at train time therefore specific light positions had to be assigned to the networks to be trained for the dataset. [18] performs best with a priori initialisation of the specularity parameter c (0 is fully specular, 1 fully Lambertian) we used 0.1 for the Cup, 0.2 for the Bell, 0.25 for the Bawl and Tool, 0.5 for the Ball, Die, Hippo, Jar and Squirrel, 0.75 for the Bunny and 1 for the rest. For [32], we used the Cauchy estimator with 0.5 on the respective hyper-parameter. For both [18] and [32] we disabled the lighting calibration parameter. For all methods, we evaluated on full resolution images (2048x1536) except for [35], which is severely limited by GPU RAM so we had to subsample to (512x384) which was the maximum we could fit on 24GB Nvidia Titan RTX. All other approaches are CPU RAM limited but 'only' require around 120GB. The computation time was varied from around 15 minutes (the fastest was [18] on the Bowl) to around two hours (the slowest was [20] on the Jar). For all of the methods, the initialisation was a flat plane at the mean depth computed exactly using the GT depth map.

Finally, we also evaluated the far-field method [13]. The assumed lighting direction was set the average one for each light and numerical integration was used on the output normal map to be able to compare surfaces. It is worth to mention that our 52 lights is within the range of lights the model in [13] is trained for.

Evaluation metrics. As it is the standard in PS literature, we first evaluate the competing approaches using the angular error on normal maps. We note that [18] and [32] output surfaces as dense depth maps, therefore the normals have been estimated using first order (forward and downward) finite differences. The other 3 methods output both surfaces and normals. It is very important to mention, that normal evaluation has two major limitations. Firstly, for real data, there can be regions where the ground truth normal uncertainty is non-negligible. This is inevitable due to capturing surfaces with a laser scanner that only provides very dense point clouds. Even micro-meter accuracy on the surface can generate a few degress uncertainty of normals in regions of complicated geometry. The second important issue with evaluating on normals is that even on synthetic data, ground truth normals are not fully consistent with the ground truth depth [33, 42]. This is inevitable due to the fact that for any non-trivial object, the projection operation generates a depth map that is discontinuous and non-differentiable for a significant portion of the pixels. In fact, to quantify this discon-

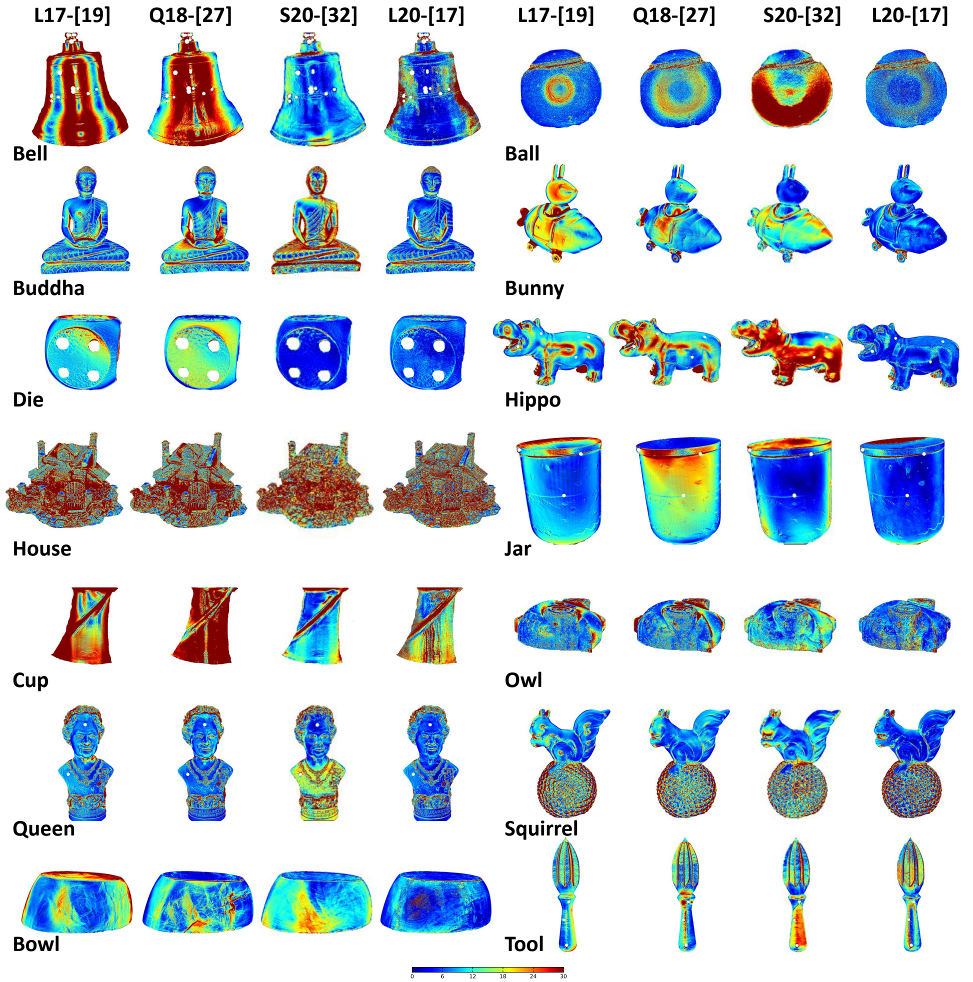

Figure 4: Normal error map comparison for all objects and all near-field methods.

tinuity measure, we compared the ground truth normals with the normals that are obtained with differentiation of the ground truth depth and indeed observed a 3.3^o error on average over the whole dataset (varying from 1.8^o on the Jar to 9.2^o on the House, as shown on the penultimate row of Table 1). Conversely, numerical integration (using [28]) of GT normals has an average error of 1.67mm with repsect to the ground truth.

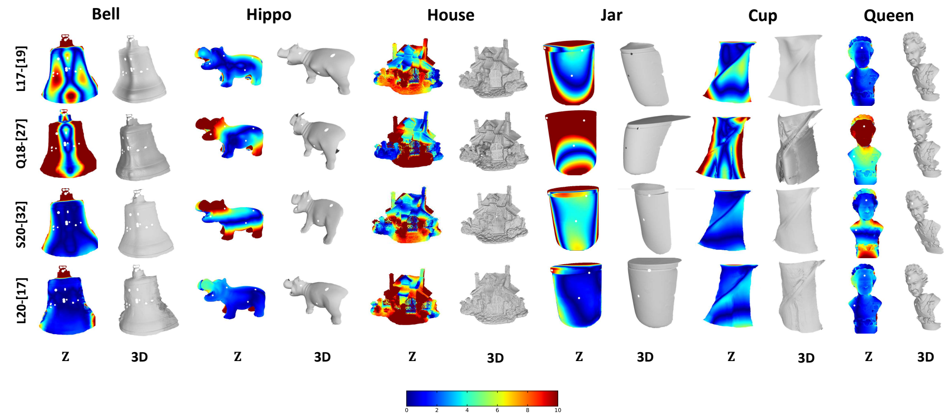

Figure 5: Output surface comparison for 6 objects and all methods. This is shown qualitatively through the 3D meshes and well as depth Z error maps (errors in mm).

5 Results

In this section, we analyse the performance of [13, 18, 20, 32, 35] which is shown quantitatively at Table 1 and qualitatively at Figures 4 and 5.

We first of all observe that the far-field method [13] fails to produce accurate results as expected. The two classical optimisation methods [18, 32] are outperformed by the deep learning approaches [35] and [20] on the normal error metric for most scenarios although [18] is not significantly worse for a few objects and indeed achieves the minimum depth error on the Queen and Tool objects. Despite the fact that [20] is the best overall performer, we emphasise that it requires knowledge of the setup at train time. [35] achieves the best performance on object with specular and metallic materials (Die & Cup both normals & depth, Bell normals only), because of the use of a patch-based network that extract the most information of the metallic object. It also achieves best MAE on the Bell. The orthographic camera assumption of [35] in terms of error translates to a growing inaccuracy towards the external part of the reconstruction (see Bell, Cup and Jar in Figure 4).

We also notice that the normal predictions are more noisy as opposed to depth prediction. This could be due to actual noisy estimates of the normals from ground truth meshes which is inevitable for any laser scanner (see in particular the Ball in Figure 4). As the ground truth depth is more reliable, it is a better evaluation metric compared to the 'ground truth' normals. See Figure 5 for depth evaluation.

By looking at Figure 4 it can be seen that even the best methods perform poorly for recovering the geometry of very oblique regions. This is observable for the Jar, Owl and Cup in Figure 4. Therefore these represent quite hard regions to retrieve.

An interesting observation is that for both CNN-based methods [20, 35], the material's specularity does not seem be a significant factor of performance. Indeed, convex regions (where self reflections are negligible) are consistently recovered correctly regardless of the material: diffuse head of Queen, bronze Bell, plastic Hippo, wooden Bowl; with the only exception being the aluminium Cup. This is a clear advantage of CNN methods against the classical ones that require diffuse or mostly diffuse materials.

We observe that the hardest regions are the ones containing high frequency details (sharp boundaries) such as House, bottom part of the Squirrel, details of the Queen etc.

6 Conclusion

In this work we proposed the first dataset for the near-field PS problem. Differently from the far-field assumption, when the target object is close to the camera/light setup several non-linear physical effects as anisotropic light propagation, light attenuation and perspective viewing geometry occur. Current PS datasets, mostly consider scenarios where such effects are negligible as they provide directional light coordinates and use quite long-focus lenses giving orthographic viewing geometry.

Recent research trends on 3D reconstruction using PS have shown an increasing interest to deal under near-field settings. However, the lack of a dataset for this topic has prevented to fairly compare different approaches. For this reason, we benchmark the recent near-field PS approaches and analyse their performance over our dataset which include objects with a wide variety of materials. We also provide a discussion about appropriate ways of evaluation (depth vs normals). In addition, as we noticed that most of the error is expectedly concentrated on the edges and discontinuity regions we conclude that future research has to improve the interpretation of the PS imaging data in these specific areas and possibly exploiting networks with edge detection capability to better deal with interreflctions.

Finally, it is worth investigating the possibility of using completely raw image data without demosicing or radial distortion compensation. This requires incorporating the radial distortion into the image irradiance equation and treating the images as pure intensity and ignoring the potential of recovering colours. The advantage of skipping these two pre-processing steps is the potential of eliminating some image artefacts, especially around image edges, which currently achieve the least accuracy.

References

[1] H. Aanæs, A. Lindbjerg Dahl, and K. S. Pedersen. Interesting interest points - A comparative study of interest point performance on a unique data set. *Int. J. Comput. Vis.*, 97(1):18–35, 2012.

[2] H. Aanæs, R. Ramsbøl Jensen, G. Vogiatzis, E. Tola, and A. B. Dahl. Large-scale data for multiple-view stereopsis. *Int. J. Comput. Vis.*, 120(2):153–168, 2016.

[3] J. Ackermann and M. Goesele. A survey of photometric stereo techniques. *Found. Trends Comput. Graph. Vis.*, 9(3-4):149–254, 2015.

[4] N. Alldrin, T. E. Zickler, and D. J. Kriegman. Photometric stereo with non-parametric and spatially-varying reflectance. In *CVPR*, 2008.

[5] M. Chandraker, J. Bai, and R. Ramamoorthi. On Differential Photometric Reconstruction for Unknown, Isotropic BRDFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 35(12):2941–2955, 2013.

[6] G. Chen, K. Han, and K.-Y. K. Wong. Ps-fcn: A flexible learning framework for photometric stereo. In *15th European Conference on Computer Vision (ECCV)*, 2018.

[7] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia. Meshlab: an open-source mesh processing tool. Eurographics, 2008.

[8] J. J. Clark. Active photometric stereo. In *1992 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1992.

[9] T. Collins and A. Bartoli. 3d reconstruction in laparoscopy with close-range photometric stereo. In *MICCAI 2012*, 2012.

[10] A. P. Harrison and D. Joseph. Maximum likelihood estimation of depth maps using photometric stereo. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(7):1368–1380, 2012.

[11] S. Herbort and C. Wöhler. An introduction to image-based 3d surface reconstruction and a survey of photometric stereo methods. *3D Research*, 2(4), 2011.

[12] T. Higo, Y. Matsushita, N. Joshi, and K. Ikeuchi. A hand-held photometric stereo camera for 3-D modeling. In *2009 IEEE International Conference on Computer Vision (ICCV)*, pages 1234–1241, 2009.

[13] S. Ikehata. Cnn-ps: Cnn-based photometric stereo for general non-convex surfaces. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 3–18, 2018.

[14] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Robust photometric stereo using sparse regression. In *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[15] K. Ikeuchi. Determining surface orientations of specular surfaces by using the photometric stereo method. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 3(6):661–669, 1981.

[16] Y. Iwahori, H. Sugie, and N. Ishii. Reconstructing shape from shading images under point light source illumination. In *International Conference on Pattern Recognition (ICPR)*, pages 83–87, 1990.

[17] Michael Kazhdan and Hugues Hoppe. Screened poisson surface reconstruction. *ACM Transactions on Graphics (ToG)*, 32(3):1–13, 2013.

[18] F. Logothetis, R. Mecca, and R. Cipolla. Semi-calibrated near field photometric stereo. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 3, page 8, 2017.

[19] F. Logothetis, R. Mecca, and R. Cipolla. A differential volumetric approach to multi-view photometric stereo. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1052–1061, 2019.

[20] F. Logothetis, I. Budvytis, R. Mecca, and R. Cipolla. A cnn based approach for the near-field photometric stereo problem. *BMVC*, 2020.

[21] F. Logothetis, I. Budvytis, R. Mecca, and R. Cipolla. Px-net: Simple, efficient pixel-wise training of photometric stereo networks. *arXiv preprint arXiv:2008.04933*, 2020.

[22] W. Matusik, H. Pfister, M. Brand, and L. McMillan. A data-driven reflectance model. *ACM Transactions on Graphics*, 22(3):759–769, 2003.

[23] R. Mecca, A. Tankus, A. Wetzler, and A. Bruckstein. A direct differential approach to photometric stereo with perspective viewing. *SIAM Journal on Imaging Sciences*, 7(2): 579–612, 2014. doi: 10.1137/120902458.

[24] R. Mecca, A. Wetzler, A. Bruckstein, and R. Kimmel. Near Field Photometric Stereo with Point Light Sources. *SIAM Journal on Imaging Sciences*, 7(4):2732–2770, 2014.

[25] A. Ngan, F. Durand, and W. Matusik. Experimental analysis of BRDF models. In *Proceedings of the Sixteenth Eurographics conference on Rendering Techniques*, pages 117–126, 2005.

[26] Y. Nie and Z. Song. A novel photometric stereo method with nonisotropic point light sources. In *ICPR 2016*, pages 1737–1742. IEEE, 2016.

[27] T. Papadhimitri and P. Favaro. Uncalibrated Near-Light Photometric Stereo. In *British Machine Vision Conference (BMVC)*, 2014.

[28] Y. Quéau and J.-D. Durou. Edge-preserving integration of a normal field: Weighted least squares, TV and L1 approaches. In *SSVM*, 2015.

[29] Y. Quéau, F. Lauze, and J.-D. Durou. Solving uncalibrated photometric stereo using total variation. *J. Math. Imaging Vis.*, 52(1):87–107, 2015.

[30] Y. Quéau, R. Mecca, and J.-D. Durou. Unbiased photometric stereo for colored surfaces: A variational approach. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4359–4368, 2016.

[31] Y. Quéau, T. Wu, and D. Cremers. Semi-calibrated near-light photometric stereo. In *International Conference on Scale Space and Variational Methods in Computer Vision (SSVM)*, pages 656–668, 2017.

[32] Y. Quéau, B. Durix, Tao Wu, D. Cremers, F. Lauze, and J.-D. Durou. Led-based photometric stereo: Modeling, calibration and numerical solution. *Journal of Mathematical Imaging and Vision (JMIV)*, 60(3):313–340, 2018.

[33] Y. Quéau, J.-D. Durou, and J.-F.çois Aujol. Normal integration: A survey. *J. Math. Imaging Vis.*, 2018.

[34] H. Santo, M. Waechter, W.-Y. Lin, Y. Sugano, and Y. Matsushita. Light structure from pin motion: Geometric point light source calibration. *Int. J. Comput. Vis.*, 2020.

[35] H. Santo, M. Waechter, and Y. Matsushita. Deep near-light photometric stereo for spatially varying reflectances. In *European Conference on Computer Vision (ECCV)*, 2020.

[36] B. Shi, Z. Mo, Z. Wu, D. Duan, S. K. Yeung, and P. Tan. A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018.

[37] Boxin Shi, Zhe Wu, Zhipeng Mo, Dinglong Duan, Sai-Kit Yeung, and Ping Tan. A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3707–3716, 2016.

[38] A. Tankus and N. Kiryati. Photometric stereo under perspective projection. In *2005 IEEE International Conference on Computer Vision (ICCV)*, pages 611–616, 2005.

[39] L. B Wolff. Relative brightness of specular and diffuse reflection. *Optical Engineering*, 33(1):285–294, 1994.

[40] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):134–144, 1980.

[41] Y. Xiong, A. Chakrabarti, R. Basri, S. J. Gortler, D. W. Jacobs, and T. E. Zickler. From shading to local shape. *IEEE Trans. Pattern Anal. Mach. Intell.*, 37(1):67–79, 2015.

[42] D. Zhu and William A. P. S. Least squares surface reconstruction on arbitrary domains. In *Proc. of the European Conference on Computer Vision (ECCV)*, 2020.