UNCALIBRATED AND UNSUPERVISED PHOTOMETRIC STEREO WITH PIECEWISE REGULARIZER

Alejandro Casanova and Antonio Agudo

Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Spain

ABSTRACT

Photometric stereo is a technique for recovering a rigid object's 3D shape, reflectance properties, lighting conditions, and specular highlights from multiple images captured under varying lighting conditions. Variational, uncalibrated, and unsupervised formulations have recently provided detailed and robust solutions to the problem, reducing the need for prior knowledge about shape geometry or lighting conditions. However, uncalibrated methods, especially when applied to real-world data, may be susceptible to noise and depth errors near boundaries or self-occlusions, stemming from missing or noisy data, surface orientation ambiguity, and calibration issues. In this paper, we introduce a novel piecewise depth regularizer to mitigate these errors, enhancing stability and improving robustness against initialization errors. We demonstrate the effectiveness of our approach through evaluations on both synthetic and real-world data, showcasing its promise in enhancing the accuracy and reliability of photometric stereo for practical applications.

Index Terms— Uncalibrated Photometric Stereo, Unsupervised Vision, Piecewise Regularizer, Specular Materials.

1. INTRODUCTION

The aim of Photometric Stereo (PS) is the recovery of a rigid object's 3D shape, reflective properties, lighting characteristics, and specular highlights, from multiple visual signals captured from the same viewpoint but under varying lighting conditions [1, 2, 3, 4]. Initially, it relied on specific assumptions about lighting control, limiting its use to controlled settings with precise lighting calibration. Uncalibrated methods emerged as an alternative, but comprised an ill-posed problem due to the linear ambiguity intrinsic to the recovered normal map [5]. Efforts to address those ambiguities introduced variational formulations, representing 3D reconstruction as a depth map rather than normal vectors [6, 7], and where later extended to general lighting conditions [8, 9, 10, 11], and non-Lambertian materials [11]. However, their overall quality notably depended on the initialization of the depth map [12, 13, 14]. Deep-learning approaches have also been

applied to PS [15, 16, 17], utilizing ground-truth parameters for supervision during training. However, these methods lack a clear physical interpretation, hindering the understanding of the interactions between lighting, 3D geometry, and specular highlights. Furthermore, they require extensive training data, which can be hard and expensive to obtain in practice.

It has been noted in recent works [10, 11, 18] that uncalibrated methods often present various forms of depth errors on their recovered shapes, especially near object boundaries and self-occlusions. These regions are particularly susceptible to inaccuracies due to various factors. Firstly, partial information near boundaries can result in missing or ambiguous data, making accurate estimation challenging. Ambiguity arises because contours mark transitions between different surface orientations, making it difficult for the algorithm to distinguish between these orientations, particularly under less-than-ideal lighting conditions or when specularities, shadows and noise are present. Boundary reflections add another layer of complexity, as objects near boundaries may reflect light differently due to their proximity to background or neighbouring objects, introducing noise, and further affecting accuracy. Moreover, calibration issues, especially pronounced near boundaries, can lead to depth errors as well, as even minor calibration errors can result in significant inaccuracies. One effective solution to mitigate depth errors is the application of regularization techniques [19, 20, 21, 22], helping to smooth out noisy or ambiguous areas. Exploring advanced PS algorithms designed to handle complex surface properties, shadows, specularities, and boundary effects can also be beneficial.

To address the existing limitations of current methods, we present an unsupervised, unified, and uncalibrated PS algorithm, capable of functioning effectively under diverse lighting conditions, and catering to a wide spectrum of objects and materials with no extra sensors, training data, or any form of ground truth information. Our method utilizes a convex formulation for shape initialization from a single visual signal, which excels in fine detail recovery. Furthermore, we introduce a physically-aware piecewise depth regularizer, to separately target depth errors in the shape boundary and enforce smoothness on the rest of the recovered surface, thus making our algorithm more robust against noise and errors stemming from calibration or initialization.

---

2. PIECEWISE PHYSICALLY-AWARE MODEL

Consider $\{\mathcal{I}_c^i \subset \mathbb{R}^2\}$ as a collection of $i = \{1, \ldots, I\}$ visual signals captured under varying illumination conditions, where $c = \{1, \ldots, C\}$ denotes different color channels. These signals encompass a rigid object intended for reconstruction. In the context of this object, we define $\mathcal{S} \subset \mathcal{I}_c^i$, representing the shape segmentation within the signal set, and $\mathcal{B} \subset \mathcal{S}$ the silhouette boundary of this object, comprising the set of pixels residing solely on the boundary of the object's silhouette. Additionally, we define $\mathcal{W} = \mathcal{S} \setminus \mathcal{B}$ as the area enclosed by the silhouette boundary. Following the principles of the Phong reflection model [23], the surface reflectance across all P pixel points $\mathbf{p} = [u, v] \in \mathcal{S}$ can be mathematically represented by aggregating elementary luminance contributions –in a unit sphere in $\mathbb{R}^3$, denoted by $\mathbb{S}^2$– originating from all incident lighting directions $\boldsymbol{\omega}$ as:

$$\mathcal{I}_c^i(\mathbf{p}) = \int_{\mathbb{S}^2} \rho_c(\mathbf{p})\, l_c^i(\boldsymbol{\omega}) \max\{0, \boldsymbol{\omega} \cdot \mathbf{n}(\mathbf{p})\}\, d\boldsymbol{\omega} + s^i(\mathbf{p}), \ (1)$$

where $\rho_c(\mathbf{p}) \in \mathbb{R}^+$ signifies the color-wise albedo, and $l_c^i(\boldsymbol{\omega})$ represents the intensity of incident lights in the respective color channels. Furthermore, $\mathbf{n}(\mathbf{p})$ corresponds to the unit-length surface normal at the surface point corresponding to the p-th pixel. Additionally, $s^i(\mathbf{p})$ characterizes the specular reflection component, and $\cdot$ indicates a dot product.

Given these components, the PS problem in an uncalibrated context involves the recovery of the 3D shape of the object (characterized by its normal vectors $\mathbf{n}(\mathbf{p})$) along with the parameters $\{\rho_c\}$, $\{l_c^i\}$ and $\{s^i\}$, all of which are derived from the input visual signals $\{\mathcal{I}_c^i\}$. To make this problem treatable, numerous studies model the irradiance map using spherical harmonics for general lighting conditions [24], often considering first or second-order approximations. Following said approximation, the image formation model described in Eq. (1) can now be expressed as follows:

$$\mathcal{I}_c^i(\mathbf{p}) \approx \rho_c(\mathbf{p})\, \mathbf{l}_c^i \cdot \mathbf{h}[\mathbf{n}](\mathbf{p}) + s^i(\mathbf{p}). \tag{2}$$

being $\mathbf{h}[\mathbf{n}] \in \mathbb{R}^\tau$ and $\mathbf{l}_c^i \in \mathbb{R}^\tau$ the first- or second-order harmonic lighting images and coefficients, with $\tau \in \{4, 9\}$, respectively. To bypass the complexity associated with inferring normal vectors in a nonlinear manner, each surface normal $\mathbf{n}[z]$ is parameterized based on its depth under perspective projection [10, 18].

Estimation Error Formulation. The model parameters can be estimated through the minimization of the image rendering error computed across all observed points within the entire set of visual signals. The residual function $r_{c,p}^i$ represents the difference between the anticipated intensity and the actual one at the p-th pixel, expressed as:

$$r_{c,p}^i(\alpha_p, \rho_{c,p}, \mathbf{l}_c^i, s_p^i, z) = \rho_{c,p}\, \mathbf{l}_c^i \cdot \mathbf{h}_p[\bar{\mathbf{n}}_p[z]/\alpha_p] + s_p^i - \mathcal{I}_{c,p}^i,$$

where $\alpha_p = |\bar{\mathbf{n}}_p[z]| \ \forall p \in \{1, \ldots, P\}$ and $\mathbf{n}_p[z] = \bar{\mathbf{n}}_p[z]/\alpha_p$, being $\bar{\mathbf{n}}_p[z]$ a linear parametrization on depth. In addition to the data term, four regularization priors are introduced. The first two enforce smoothness and sparsity on the albedo and specular maps, respectively. The last two represent the piecewise regularization terms of the depth map. The former enforces smoothness on the object's boundary, whereas the latter targets the whole surface within. Then, the total cost function $\mathcal{D}(\alpha, \{\rho_c\}, \{\mathbf{l}_c^i\}, \{s^i\}, z)$ can be written as:

$$\sum_{i=1}^{I} \sum_{c=1}^{C} \sum_{p=1}^{P} \varphi_\lambda(r_{c,p}^i(\alpha_p, \rho_{c,p}, \mathbf{l}_c^i, s_p^i, z_p))$$

$$+ \mu_a \sum_{c=1}^{C} \sum_{p=1}^{P} |(\nabla \rho_c)_p|_\gamma + \mu_s \sum_{i=1}^{I} \sum_{p=1}^{P} |(s^i)_p|_{\gamma_s}$$

$$+ \mu_b \sum_{b=1}^{B} |(\nabla z)_b|_{\gamma_b} + \mu_w \sum_{w=1}^{W} |(\nabla z)_w|_{\gamma_w}, \tag{3}$$

where μ_a, μ_s, μ_b and μ_w are the weight parameters for each regularization term, $|\cdot|_\gamma$ signifies a Huber norm, ∇ is the spatial gradient operator, and $\varphi_\lambda(x) = \lambda^2 \log(1 + \frac{x^2}{\lambda^2})$ is Cauchy's M-estimator, where λ is a scaling coefficient.

Given that the previous problem is highly non-linear and non-convex, a proper initialization is essential to achieve accurate results. For this reason, we implemented the minimal surface initialization proposed in [12], which is capable of inferring surfaces from a single visual signal by considering a volume initialization encoded by a factor κ that allows the monocular reconstruction to be scaled, such that:

$$\sum_{p=1}^{P} \sqrt{1 + |(\nabla z)_p|^2} + \upsilon \sum_{p=1}^{P} z_p - \kappa, \tag{4}$$

where υ is a weight coefficient. Dirichlet's boundary conditions are also enforced as $z_p = 0, \ \forall p \in \mathcal{B}$. Problem in Eq. (4) is solved by gradient descent. Ultimately, a lagged block coordinate descent algorithm is used to minimize $\mathcal{D}$ in Eq. (3).

3. EXPERIMENTAL EVALUATION

We now provide both quantitative and qualitative evaluations on the performance of our method, on both synthetic and real-world data. For this purpose we compute the mean angular error between the estimated $\mathbf{n}[z]$ and ground-truth $\mathbf{n}^{gt}[z]$ normal vectors as:

$$\text{MAE} = \frac{1}{P} \sum_{p=1}^{P} \cos^{-1}\left(\frac{\mathbf{n}_p^{gt}[z] \cdot \mathbf{n}_p[z]}{|\mathbf{n}_p^{gt}[z]| \cdot |\mathbf{n}_p[z]|} \right). \tag{5}$$

However, as it will be shown later, this metric can prove insufficient when surface discontinuities, like self-occlusions,

Dataset / Meth.	Armadillo		Joyful Yell		Lucy		Thai		Average	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
UPS [10]	15.76	22.28	10.41	0.42	27.70	214.26	15.29	30.73	17.29	66.92
USPS [11]	**14.56**	24.58	**7.93**	0.55	16.02	172.48	15.29	28.04	13.45	56.41
Ours	14.58	**14.91**	8.84	**0.36**	**15.17**	**92.58**	**8.86**	**25.19**	**11.86**	**33.26**
Relative error w.r.t. [10]	1.08	1.49	1.18	1.17	1.82	2.31	1.73	1.22	1.46	2.01
w.r.t. [11]	0.99	1.65	0.90	1.53	1.06	1.86	1.73	1.11	1.13	1.70

Table 1. **Quantitative evaluation on synthetic data.** The table reports the MAE in degrees and RMSE for depth, respectively, for each dataset.

are present (due to the ambiguity intrinsic to surface normal representation). For this reason, we also compute a root-mean-square error between the estimated z_p and ground-truth z_p^{gt} depth maps as:

$$\text{RMSE} = \sqrt{\frac{1}{P}\sum_{p=1}^{P}(z_p - z_p^{gt})^2}. \qquad (6)$$

We propose using both metrics in a complementary manner, in order to achieve a thorougher evaluation. Additionally, both errors can be plotted (before being averaged) to serve a more qualitative evaluation of the results. The angular error metric better serves to quantitatively evaluate the overall quality of the reconstruction, and can be used to locate and visualize errors of a more subtle and localized nature. On the other hand, the depth error metric can serve to identify and evaluate errors that extend over a larger scale, such as smooth deformations that extend across large surfaces, or offsets caused by surface discontinuities.

Synthetic datasets. We considered four challenging synthetic shapes under varying light conditions: *Joyful Yell* [25], *Armadillo*, *Lucy*, and *ThaiStatue* [26], which were generated following [11]. Firstly, we tuned the depth initialization algorithm [10], and found that the optimal κ values for our implementation were 32, 3, 5 and 15 for *Joyful Yell*, *ThaiStatue*, *Armadillo* and *Lucy*, respectively. Secondly, we proceeded with the tuning of the weight parameters in $\mathcal{D}(\cdot)$, and fixed the values that best helped minimize said cost function. For our implementation, these values were $\gamma = \gamma_s = 0.1$, $\mu_a = 5 \cdot 10^{-3}$, $\mu_s = 2 \cdot 10^{-6}$, $\mu_b = 5 \cdot 10^{-7}$, $\mu_w = 5 \cdot 10^{-10}$ and $\lambda = 0.15$. We use the same values for all our experiments.

Our results are reported in Table 1, obtaining MAE and RMSE errors of 11.86 and 33.26 on average, respectively. To make a fair comparison, we also include the results by applying the competing approaches UPS [10] and USPS [11], as in all cases the depth initialization strategy is the same. It is worth noting that our method outperforms the solution in UPS [10] and USPS [11] by large margins in terms of both MAE (from 46% to 13% on average, respectively) and RMSE (from 101% to 70% on average, respectively) metrics. Thanks to our piecewise physically-aware formulation, the algorithm is more stable and less dependent on depth initialization (see Fig. 1). However, as our formulation exploits more sophisticated constraints to solve the problem, it is slightly more
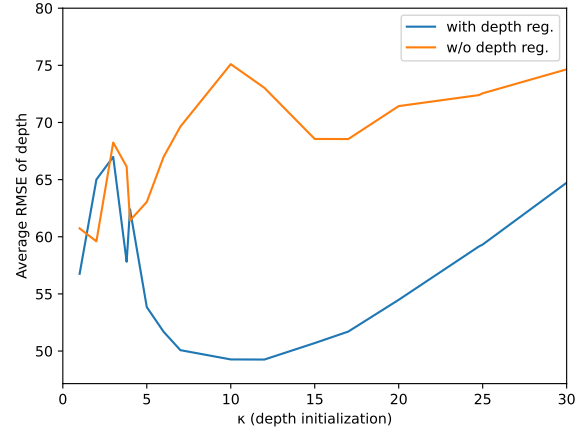

Fig. 1. **Evaluation on the impact of depth initialization.** Average RMSE of depth across all four synthetic datasets, for different κ values, with and without depth regularization. κ is the hyper-parameter that determines the volume of the balloon-like shape to be generated by the initialization algorithm, we use in Eq. (4).

Method / Dataset	Armadillo	Joyful Yell	Lucy	Thai	Average
USPS [11]	**107.70**	**148.84**	**194.69**	**261.87**	**178.28**
Ours	156.08	217.71	259.61	363.81	249.30
Speedup w.r.t. [11]	0.69	0.68	0.75	0.72	0.72

Table 2. **Quantitative evaluation of execution times for the synthetic dataset.** The table reports the execution times in seconds and speedup for each dataset.

expensive. Particularly, for a non-optimized Python implementation, our method increases the computing time by approximately 30% (see Table 2). In any case, our algorithm provides a competitive trade-off between accuracy and computational cost with respect to state-of-the-art approaches.

Figure 2 shows an example of how both the MAE and RMSE metrics were used in a complementary manner to evaluate unfavourable results. The recovered shape presents errors in texture and detail, especially on the face, chest and base of the statue, which are better appreciated in the normal map and angular error representations. Furthermore, self-

Fig. 2. **Shape recovery evaluation.** From left to right: ground-truth and recovered normal maps, ground truth and recovered shape, angular error map, and depth squared error map.
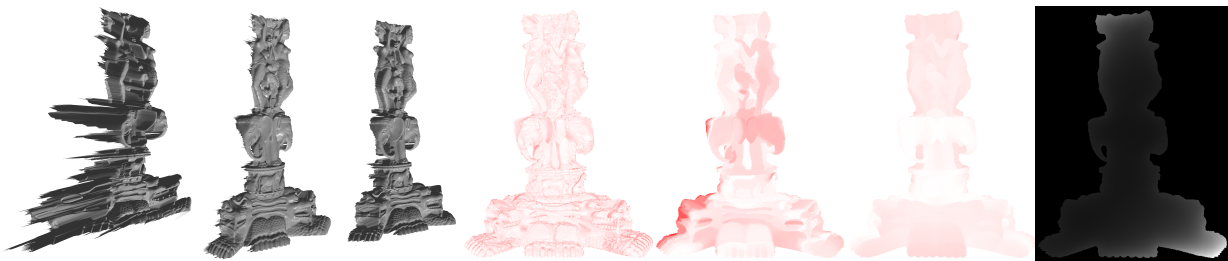

Fig. 3. **Qualitative evaluation of the piecewise depth regularizer.** From left to right: shape recovered with unfavourable tuning and no regularizer, shape recovered with the same tuning but using depth regularization, ground-truth shape, angular error before regularization, depth error before regularization, depth error after regularization, exaggerated representation of the shape used for depth initialization.

occlusions present on the original shape (i.e., the folds of the dress or the gap between the wings and the rest of the body) are lost after reconstruction, and the recovered shape is somewhat flattened. This difference can only be appreciated by comparing the rendered shapes or by plotting the depth error metric, which directly targets the depth maps, instead of the surface normals.

Furthermore, Fig. 3 shows the rather unfavourable result obtained for the *Thai* dataset when certain trade-off values where fixed during tuning. Although the angular error map does not provide any particular insights on the quality of the reconstruction, the depth error map does show that the recovered shape presents considerable noise on the left boundary of the base, and is warped along the diagonal axis (see Fig. 3-fifth column). Taking into account how critically dependent our algorithm is on depth initialization [10], we were able to trace this deformation back to the shape generated by the depth initialization algorithm, which also presented a very subtle deformation of the same nature (see Fig. 3-seventh column), and which propagated to the final result. The depth regularizer addressed initialization errors as well as the issue of noise, showcasing a more robust response across all datasets (see Fig. 1).

Real datasets. We now evaluate our algorithm on eight real-world visual collections. The first four datasets are *Cat*,

Dataset / Meth.	Cat	Frog	Hippo	Scholar	Average
StLS [27]	14.83	**11.80**	20.25	28.13	18.75
UPS [10]	9.53	16.01	**11.22**	12.93	12.42
USPS [11]	9.65	14.83	11.59	12.95	12.26
Ours	**9.47**	12.10	**11.22**	**12.72**	**11.38**
Relative error w.r.t. [27]	1.57	0.98	1.80	2.21	1.65
w.r.t. [10]	1.01	1.32	1.00	1.02	1.09
w.r.t. [11]	1.02	1.23	1.30	1.02	1.08

Table 3. **3D reconstruction evaluation and comparison.** The table reports MAE results in degrees for StLS [27], UPS [10], USPS [11] and our algorithm.

Frog, *Hippo*, and *Scholar* [27], which present varying geometries, diffuse albedo, and were captured under directional lighting. Due to the nature of the lighting conditions, these shapes present strong shadow regions, caused by the interaction between self-occlusions and the directional light rays. Table 3 shows a quantitative evaluation between our results and those obtained by the competing approaches StLS [27], UPS [10] and USPS [11]. As it can be seen, our method obtains the best solution, obtaining between the 8% and 65% of performance on average. In Fig. 4 we present our joint estimation results, which are physically plausible and detailed, thus proving that our algorithm is generalizable to mainly di-
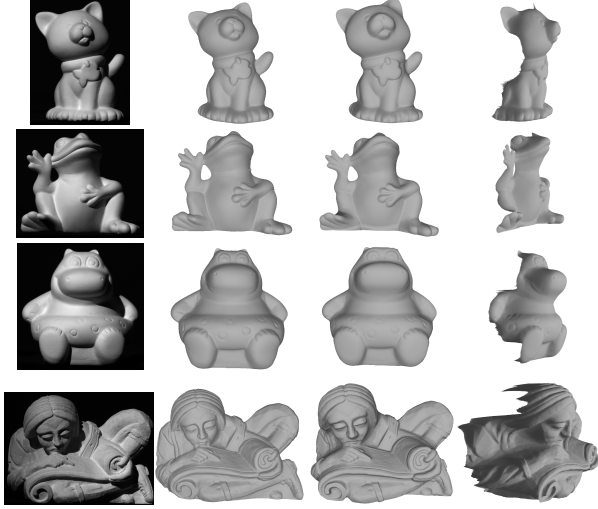
Fig. 4. **Qualitative evaluation on real datasets.** From top to bottom: *Cat*, *Frog*, *Hippo*, and *Scholar*. From left to right: an arbitrary input image, the 3D ground truth, our 3D estimation and a novel point of view of that shape.

Fig. 5. **Qualitative evaluation on real datasets**. From top to bottom: *Vase*, *Face1*, *Face2*, and *Tablet*. First column shows an arbitrary input image. Columns 2 & 4 represent estimation by [18], for two points of view. Columns 3 & 5 display the same information for our method.

rectional lighting conditions, and provide satisfying results in spite of the presence of strong self-cast shadows.

The last four shapes we evaluated present a wide variety of natural geometries, highly detailed surfaces, and different examples of natural albedos. The visual collections are *Face1*, *Face2*, *Tablet*, and *Vase* [28], which were captured under daylight and a freely moving LED. As shown in Fig. 5, our estimations are physically plausible and consistent with the input images. The previous attempt on this same dataset [18] presented noisy regions in the *Face1* and *Face2* scenarios, as well as over-deformations in the *Vase* collection. Our implementation addressed these issues thanks to the use of piecewise depth constraints. Particularly noticeable is the improvement on the *Face1* dataset, where the tip of the nose presented considerable noise due to specular reflections. Furthermore, on the *Tablet* case our algorithm did not perform as well. The recovered shape already presented overall smoothness, and depth regularization led to a loss in detail (the groove on the case is shallower in our case). A certain loss of surface texture or detail is an inherent trade-off of depth regularization, which otherwise effectively helps to tackle noise and initialization problems. Furthermore, the recovery of planar shapes (such as that of the tablet case) is especially challenging through monocular vision, which explains the quality of this particular reconstruction.

4. CONCLUSION

In this study, we have proposed an uncalibrated, unified, and unsupervised PS algorithm that enhances performance, stability, and robustness across diverse datasets. A piecewise physically-aware regularizer was proposed to address
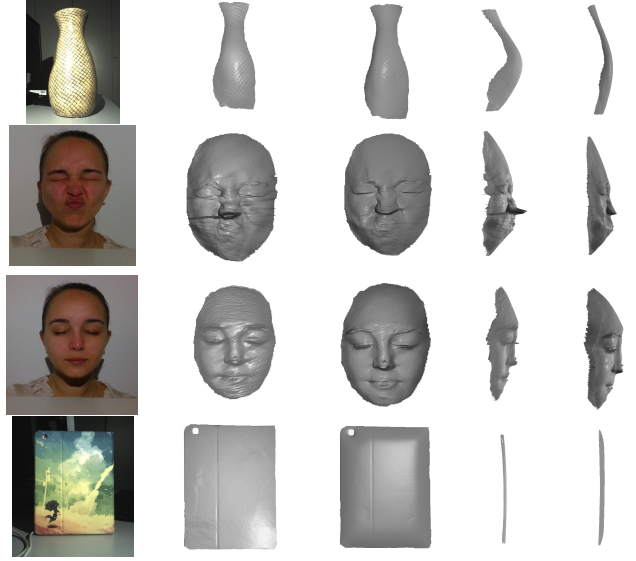
depth errors by targeting object boundaries and smoothing out noisy regions, which significantly improved stability. We also tackled the issue of inaccurate depth initialization, which can cause shape deformations. Our method demonstrated resilience against initialization errors, resulting in more accurate outcomes. Nevertheless, it is essential to note that depth regularization entails a trade-off, as excessive application can lead to a loss in detail and flattened geometry. Therefore, a balanced approach is necessary when employing this technique. Our future work aims to take our formulation a step further by addressing outdoor scenarios with strong occlusions.

5. REFERENCES

[1] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *OE*, vol. 19, no. 1, pp. 191139, 1980.

[2] T. P. Wu and C. T. Tang, "Photometric stereo via expectation maximization," *TPAMI*, vol. 32, no. 3, pp. 546–560, 2007.

[3] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma, "Robust photometric stereo via low-rank matrix completion and recovery," in *ACCV*, 2010.

[4] S. Ikehata, "CNN-PS: CNN-based photometric stereo for general non-convex surfaces," in *ECCV*, 2018.

[5] H. Hayakawa, "Photometric stereo under a light source with arbitrary motion," *JOSAA*, vol. 11, no. 11, pp. 3079–3089, 1994.

[6] M. Chandraker, J. Bai, and R. Ramamoorthi, "On differential photometric reconstruction for unknown, isotropic BRDFs," *TPAMI*, vol. 35, no. 12, pp. 2941–2955, 2013.

[7] Y. Queau, T. Wu, F. Lauze, J. D. Durou, and D. Cremers, "A non-convex variational approach to photometric stereo under inaccurate lighting," in *CVPR*, 2017.

[8] S. Peng, B. Haefner, Y. Queau, and D. Cremers, "Depth super-resolution meets uncalibrated photometric stereo," in *ICCVW*, 2017.

[9] Z. Mo, B. Shi, F. Lu, S. K. Yeung, and Y. Matsushitae, "Uncalibrated photometric stereo under natural illumination," in *CVPR*, 2018.

[10] B. Haefner, Z. Ye, M. Gao, T. Wu, Y. Quéau, and D. Cremers, "Variational uncalibrated photometric stereo under general lightings," in *ICCV*, 2019.

[11] P. Estevez and A. Agudo, "Uncalibrated, unified and unsupervised specular-aware photometric stereo," in *ICPRW*, 2022.

[12] M. R. Oswald, E. Toeppe, and D. Cremers, "Fast and globally optimal single view reconstruction of curved objects," in *CVPR*, 2012.

[13] S. Vicente and L. Agapito, "Balloon shapes: reconstructing and deforming objects with volume from images," in *3DV*, 2013.

[14] A. Agudo, "Safari from visual signals: Recovering volumetric 3D shapes," in *ICASSP*, 2022.

[15] Z. Yao, K. Li, Y. Fu, H. Hu, and B. Shi, "GPS-net: Graph-based photometric stereo network," in *NeurIPS*, 2020.

[16] B. Kaya, S. Kumar, C. Oliveira, V. Ferrari, and L. Van Gool, "Uncalibrated neural inverse rendering for photometric stereo of general surfaces," in *CVPR*, 2021.

[17] B. Kaya, S. Kumar, C. Oliveira, V. Ferrari, and L. Van Gool, "Uncertainty-aware deep multi-view photometric stereo," in *CVPR*, 2022.

[18] A. Agudo, "Detail-aware uncalibrated photometric stereo," in *ICASSP*, 2023.

[19] M. Harker and P. O'Leary, "Regularized reconstruction of a surface from its measured gradient field," in *arXiv:1308.4292*, 2013.

[20] M. Harker and P. O'Leary, "Direct regularized surface reconstruction from gradients for industrial photometric stereo," *CI*, vol. 64, no. 9, pp. 1221–1228, 2013.

[21] Andrew J. Wagenmaker, Brian E. Moore, and Raj Rao Nadakuditi, "Robust surface reconstruction from gradients via adaptive dictionary regularization," in *ICIP*, 2017.

[22] D. Antensteiner and S. Štolc, "Regularization in higher-order photometric stereo inspection for non-lambertian reflections," in *VISIGRAPP*, 2020.

[23] S. Marschner and P. Shirley, *Fundamentals of Computer Graphics*, A K Peters/CRC Press, 2018.

[24] R. Basri, D. Jacobs, and I. Kemelmacher, "Photometric stereo with general, unknown lighting," *IJCV*, vol. 72, no. 5, pp. 239–257, 2007.

[25] The Joyful Yell, ," URL: http://www.thingiverse.com/thing:897412.

[26] M. Levoy, J. Gerth, B. Curless, and K. Pull, "The stanford 3D scanning repository," 2005.

[27] Y. Xiong, A. Chakrabarti, R. Basri, S. J. Gortler, D. W. Jacobs, and T. Zickler, "From shading to local shape," *TPAMI*, vol. 37, no. 1, pp. 67–79, 2015.

[28] B. Haefner, S. Peng, A. Verma, Y. Queau, and D. Cremers, "Photometric depth super-resolution," *TPAMI*, vol. 42, no. 10, pp. 2453–2464, 2019.