

Self-calibrated Near-light Photometric Stereo using a Geometric Proxy

Benjamin Couprie¹, François Lauze², Antoine Laurent^{1,3}
Jean Melou¹, Yvain Quéau⁴, and Jean-Denis Durou¹

¹ IRIT, UMR CNRS 5505, Toulouse, France

² DIKU, University of Copenhagen, Denmark

³ TRACES, UMR CNRS 5608, Toulouse, France

⁴ GREYC, CNRS, UNICAEN, ENSICAEN, Normandie Université, France

 benjamin.couprie@irit.fr

Abstract. We introduce a 3D reconstruction technique based on near-light photometric stereo, where lighting calibration is achieved on-the-fly from a rough geometric knowledge obtained, e.g., by structure-from-motion. From this coarse geometry, the proposed graduated optimisation setup estimates all light source parameters (position, intensity, orientation, and anisotropy factor of each source), which can then be used within a calibrated photometric stereo algorithm. A series of real-world experiments is conducted, which validates the interest of the proposed approach.

1 Introduction

Photometric stereo (PS) is a technique which provides 3D and appearance estimates of objects from a single view and multiple light sources. The last decade has seen major improvements in the technique and has made it an effective one for 3D digitisation. However, the quality of the results of *calibrated* PS remains highly dependent on the accuracy of illumination calibration. Therefore, *uncalibrated* PS, which relaxes the need for calibration, has long been identified as a promising alternative [7]. Unfortunately, assuming a linear illumination model (directional and uniform lighting), the solution can be recovered only up to a generalized bas-relief (GBR) transformation [22], therefore, an additional prior must be introduced [4]. Extensions of this framework to more general (spatially varying) linear illumination models have been proposed [6,12], yet they still heavily rely on hand-crafted priors for resolving the underlying ambiguities.

In contrast, *nonlinear* illumination models (where the incident illumination in each 3D point depends nonlinearly on the unknown geometry) are not prone to GBR-like ambiguities [13]. In addition, such models are particularly well suited for describing light-emitting diodes (LEDs), which represent common, versatile, inexpensive, and easy-to-use light sources. Thus, near-light PS represents a very practical 3D reconstruction method, for which efficient algorithmic solutions based on variational methods [14] or deep learning exist [9]. Nonetheless, careful calibration of the LED parameters (or at least some of them [10,16]) must

be carried out for such efficient methods to be employed. This seriously limits the interest of the approach, as near-light calibration remains cumbersome in practice [19,20].

Therefore, the goal of this paper is to provide an easy-to-use method for precise LED calibration from a relatively coarse geometry obtained by any method providing this information. Examples of plug-and-play packages for obtaining such a rough proxy include the structure-from-motion (SfM) / multiview stereo (MVS) softwares Meshroom [5] and Colmap [17,18], or the recent deep learning-based monocular geometry estimators [1,21]. If all the parameters of these complex light sources (position, intensity, principal direction and angular anisotropy – which can be colour-dependent) could be accurately deduced from this rough geometry, then self-calibration would become possible for near-light PS, enabling a simple and very accurate 3D reconstruction (see Figure 1).

In spirit, our method resembles the very recent one [2], where MVS is combined with PS within an uncalibrated multiview, near-light framework. Although yielding state-of-the-art results, the latter method remains computationally heavy, as it involves generic neural networks for modelling the surface geometry and reflectance, and it jointly solves MVS and PS. On the contrary, we rather opt for computational efficiency and modularity, by focusing on PS-based refinement of a geometric approximation. This way, we conserve network-free representations which allow resorting to a sequence of simple optimisation procedures and explicit models.

The recovery of light source parameters is actually conditioned by the scale of the scene, including the object and light sources. When the light source is sufficiently far away compared to the size of the object under scrutiny, then a directional luminous flux will in general be a very good approximation for the LED illumination in that case. At a somewhat shorter scale, a point light source spreading light radially with an inverse-of-squared-distance fall-off, becomes a much better estimate. At an even shorter scale, anisotropy plays its part: one must also account for the angular attenuation as we deviate from the principal illumination direction of the LED. Neglecting such nonlinearities in the PS model would skew the reconstructed geometry since darker patterns due to anisotropy would be falsely interpreted as some important curvature of the scene. The main other source of limitation in light source parameters recovery is the precision at which the initial geometry is known: a model that is too smooth or too noisy will also impede the estimation of sensitive parameters, such as principal direction or anisotropy. Considering these two features (scale-dependency and sensitivity) encourages us to design a graduated and robust optimisation procedure. The code can be accessed here : <https://github.com/BenjaminCouprie/LED-getter>

Contributions. In this paper, we show that a coarse to good knowledge of the geometry of an object allows us to recover important information about the LEDs and then use the full power of PS to obtain a much more precise reconstruction of the object under scrutiny. We deal with the complexity of the parameters space in several steps, from the coarse approximation of light sources as

directional to the recovery of the most important near-field anisotropic punctual sources parameters, by progressively incorporating scale information. Section 2 introduces our forward model, which we invert by introducing complexity in the model step by step in Section 3 to self-calibrate the LEDs using a geometric approximate knowledge of the scene. Our approach is evaluated in Section 4, before our conclusions are drawn in Section 5.

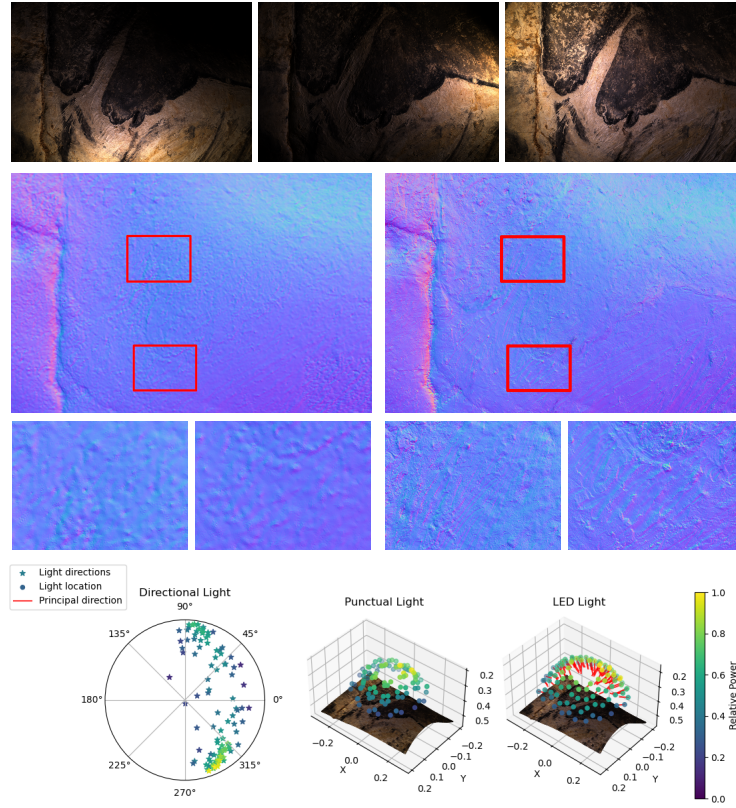


Fig. 1: Overview of the proposed framework. From a set of PS images acquired using LEDs at very short distance (top row) and a rough geometric knowledge (MVS normal map displayed in second and third rows, left), we automatically estimate all the light source parameters in a progressive manner (bottom row), enabling the subsequent use of self-calibrated PS (second and third row, right). The results of this figure will be discussed later in the article.

2 From PS to Coloured Near-light PS

Photometric Stereo. In the standard setup, N photographs of a surface $\mathcal{S}$ are taken from the same angle of view, but under varying illumination. The pictured surface is assumed to be *Lambertian* i.e., it reflects the light diffusely and the reflectance is (up to a multiplicative constant) given by the *albedo* $\rho \in [0, 1]$. When these N illuminations are directional, thus characterized by a set of vectors $\mathbf{s}^i \in \mathbb{R}^3$, $i \in \{1, \dots, N\}$, the measured brightness at pixel $\mathbf{p} = \mathbf{p}(\mathbf{x})$, which is the projection of surface point $\mathbf{x}$ onto the camera image plane, is given as:

$$I^i(\mathbf{p}) = \rho(\mathbf{x}) \{ \mathbf{n}(\mathbf{x}) \cdot \mathbf{s}^i \}_+, \quad i = 1 \dots N. \quad (2.1)$$

where $\mathbf{n}(\mathbf{x})$ is the surface normal, and $\{x\}_+ = \max\{x, 0\}$ encodes self-shadows.

Calibrated PS assumes that lighting is known, both in terms of direction and intensity. Under this assumption, with $N \geq 3$ illuminations not all coplanar, System (2.1) admits a unique solution $\mathbf{m}(\mathbf{x})$, where $\mathbf{m}(\mathbf{x}) = \rho(\mathbf{x}) \mathbf{n}(\mathbf{x})$, from which it is easy to deduce the albedo and the normal at $\mathbf{x}$. However, such a simple approach breaks in many real-world scenarios, when actual illumination cannot be fully characterized by calibrating the illumination vectors $\mathbf{s}^i \in \mathbb{R}^3$.

Near-light Models. Conventional photometric stereo assumes that the incident luminous fluxes are parallel and uniform (far away sources), which is difficult to guarantee. In such a model, illumination does not depend on the 3D point $\mathbf{x}$ on the surface, but only on a direction and an intensity:

$$\mathbf{s}^i(\mathbf{x}) = \mathbf{s}^i = \varphi^i \mathbf{t}^i, \quad (2.2)$$

with $\mathbf{t}^i \in \mathbb{S}^2$ a unit vector representing the direction of the source, and $\varphi^i > 0$ its intensity. Some works have extended the classic PS pipeline to general spatially-varying illumination models $\mathbf{s}^i(\mathbf{x})$ [12,15], yet they strongly rely on hand-crafted priors. An alternative is to resort to a parametric model making explicit the non-uniform nature of lighting (which is clearly visible in the images of Figure 1).

A very common such parametric model, encountered in the *near-field* context, is that of punctual light sources. Therein, each source is characterized by a position $\mathbf{q}^i \in \mathbb{R}^3$. It spreads light radially, and the incident illumination direction in $\mathbf{x}$ is thus explicitly given by the unit vector $\frac{\mathbf{q}^i - \mathbf{x}}{\|\mathbf{q}^i - \mathbf{x}\|}$. Besides, the source intensity φ^i is attenuated with the distance travelled according to an inverse-square fall-off law, yielding the following nonlinear illumination model:

$$\mathbf{s}^i(\mathbf{x}) = \varphi^i \frac{\mathbf{q}^i - \mathbf{x}}{\|\mathbf{q}^i - \mathbf{x}\|^3}, \quad (2.3)$$

which explicitly depends on the coordinates of $\mathbf{x}$. Since these coordinates are the unknowns of the 3D reconstruction problem, solving PS becomes much more arduous than solving a linear system as above. Nonetheless, let us stress that when $\mathbf{q}^i$ moves far away from the scene (far field), the punctual model behaves like a directional one, provided that some intensity adjustments are made.

Given their low cost and ease of use, LEDs are frequently encountered in active illumination applications. However, when used in a near-field setting, LEDs exhibit an anisotropic behaviour, with an angular intensity decay off a so-called *principal* direction [11]. A common model for this anisotropic behaviour is that of imperfect Lambertian light source, which combines point light illumination with cosine-power angular fall-off:

$$\mathbf{s}^i(\mathbf{x}) = \varphi^i \frac{(\mathbf{d}^i \cdot \boldsymbol{\sigma}^i(\mathbf{x}))^{\mu^i}}{\|\mathbf{q}^i - \mathbf{x}\|^2} \boldsymbol{\sigma}^i(\mathbf{x}), \quad (2.4)$$

where the light source is characterized by an intensity φ^i , a position $\mathbf{q}^i$ from which is deduced the incident direction $\boldsymbol{\sigma}^i(\mathbf{x}) = \frac{\mathbf{q}^i - \mathbf{x}}{\|\mathbf{q}^i - \mathbf{x}\|}$, a principal direction $\mathbf{d}^i$, and a parameter of anisotropy $\mu^i \in \mathbb{R}_+$. When $\mu^i = 0$, this anisotropic model becomes isotropic, and a perfectly Lambertian light source corresponds to $\mu^i = 1$.

Colour. So far, we have considered monochromatic sources, however real-world ones have a continuous spectrum, which affects both their intensity and their anisotropy:

$$\mathbf{s}^i(\mathbf{x}, \lambda) = \varphi^i(\lambda) \frac{(\mathbf{d}^i \cdot \boldsymbol{\sigma}^i(\mathbf{x}))^{\mu^i(\lambda)}}{\|\mathbf{q}^i - \mathbf{x}\|^2} \boldsymbol{\sigma}^i(\mathbf{x}). \quad (2.5)$$

Similarly, the scene reflectance ρ also becomes wavelength-dependent, and the proportionality coefficient absorbed in it in (2.1) becomes a wavelength-dependent sensor response $r_c : \mathcal{V} \rightarrow \mathbb{R}_+$, with $\mathcal{V}$ the visible spectrum, where the subset c refers to the colour channel (R, G and B denote the red, green and blue channels, respectively).

In the general case, the measured brightness in each channel is then obtained by integrating (2.1) over this spectrum:

$$I_c^i(\mathbf{p}) = \int_{\mathcal{V}} r_c(\lambda) \rho(\mathbf{x}, \lambda) \{\mathbf{n}(\mathbf{x}) \cdot \mathbf{s}^i(\mathbf{x}, \lambda)\}_+ d\lambda, \quad i = 1 \dots N. \quad (2.6)$$

Without loss of generality, we will consider in the following that the N LEDs are identically manufactured. It follows that they have the same parameter of anisotropy $\mu(\lambda)$, but they differ in their emission spectrum $\varphi^i(\lambda)$, their location $\mathbf{q}^i$ and their principal direction $\mathbf{d}^i$.

Assuming that the coloured filter for each channel can be assimilated to a Dirac function, i.e. $r_c = \delta_{\lambda_c}$, then (2.6) simplifies to:

$$I_c^i(\mathbf{p}) = \rho_c(\mathbf{x}) \varphi_c^i \frac{(\mathbf{d}^i \cdot \boldsymbol{\sigma}^i(\mathbf{x}))^{\mu_c}}{\|\mathbf{q}^i - \mathbf{x}\|^2} \{\mathbf{n}(\mathbf{x}) \cdot \boldsymbol{\sigma}^i(\mathbf{x})\}_+, \quad i = 1 \dots N. \quad (2.7)$$

with $\rho_c(\mathbf{x}) = \rho(\mathbf{x}, \lambda_c)$, $\varphi_c^i = \varphi^i(\lambda_c)$ and $\mu_c = \mu(\lambda_c)$.

Of course, none of the three coloured filters that characterize the Bayer pattern is perfectly monochromatic. Nevertheless, Model (2.7) seems physically realistic: the luminous flux emitted by a LED in each channel is characterized by

a maximum intensity φ_c^i in the main direction depending on the channel, as well as the anisotropy parameter μ_c . We therefore assume neither that the emission spectrum is white, nor that it is independent of the colour channel. In fact, we have observed that the apparent colour of LEDs depends on the point of view. In addition, we observed that with the LEDs we used, the light emitted in the principal direction is white. Therefore, we simplify the model by replacing φ_c^i by φ^i .

Light Sources and Scale. The above light source models involve a shift in the level of approximation from the scale of the acquisition setup. This scale, here, is the spatial extent of the measuring device, in ratio to the size of the observed object, and implies different models. Indeed, the anisotropic LED model tends to the point source model when scale increases, and the point source model tends to the directional one, when scale increases. This will be used in the next section to build an optimisation setup to recover all the light source parameters, given a coarse approximation of surface geometry.

3 A Graduated Optimisation Approach

The General Model. We have N RGB images I_c^i , $i \in \{1, \dots, N\}$, $c \in \{R, G, B\}$, from the same viewpoint, with N different light sources. As seen above, under the assumptions of Lambertian scene and identical LEDs, the image formation model is provided by:

$$I_c^i(\mathbf{p}) = \rho_c(\mathbf{x}) \{ \mathbf{n}(\mathbf{x}) \cdot \mathbf{s}^i(\mathbf{x}; \theta^i, \mu_c) \}_+ \quad (3.1)$$

where $\mathbf{s}^i(\mathbf{x}; \theta^i, \mu_c)$ is the illumination vector corresponding to (2.7). It depends on the surface point $\mathbf{x}$, as well as on a per-LED set of parameters θ^i comprising the position $\mathbf{q}^i$, the principal direction $\mathbf{d}^i$ and the intensity φ^i , and on the spectral anisotropy parameters $\boldsymbol{\mu} = (\mu_R, \mu_G, \mu_B)$ supposedly common to all sources.

We assume that we have some coarse knowledge of the geometry of the surface, represented by a discrete set of points $\mathbf{x} \in D$ over which we have the RGB measurements $I_c^i(\mathbf{p}(\mathbf{x}))$ and associated estimates of the normal $\mathbf{n}(\mathbf{x})$ and 3D coordinates $\mathbf{x}$. We wish to recover the surface colour albedo $\boldsymbol{\rho}(\mathbf{x}) = (\rho_R(\mathbf{x}), \rho_G(\mathbf{x}), \rho_B(\mathbf{x}))$ at these points, along with the source parameters $\boldsymbol{\theta} = (\theta^1, \dots, \theta^N)$ and anisotropy parameters $\boldsymbol{\mu} = (\mu_R, \mu_G, \mu_B)$, by minimizing the discrepancy between both sides of (3.1) in the sense of the classical Huber loss:

$$H_\varepsilon(a) = \begin{cases} \frac{1}{2} a^2, & \text{if } |a| < \varepsilon, \\ \varepsilon (|a| - \frac{1}{2}\varepsilon), & \text{otherwise,} \end{cases}$$

over the set D of available measurements. Our search space is therefore $\mathbb{R}^{3|D|} \times \mathbb{R}^{3N} \times \mathbb{S}^{2N} \times \mathbb{R}_+^N \times \mathbb{R}_+^3$. Here, $\mathbb{R}^{3|D|}$ is the albedo state space, $\mathbb{R}^{3N}$ is the position state space, $\mathbb{S}^{2N}$ is the principal direction state space, $\mathbb{R}_+^N$ is the intensity state space and $\mathbb{R}_+^3$ is the $\boldsymbol{\mu}$ -state space. This is a rather complex state space for a non-convex problem. To tackle it, we gradually refine the light source model, going from the coarse, directional one to anisotropic point light LEDs.

Step 1: coarse directional estimation. In that step, we assume the light sources are directional, following Model (2.2). Here each source is characterized by a vector $\mathbf{s}_0^i \in \mathbb{R}^3$, whose direction will be denoted $\mathbf{t}^i \in \mathbb{S}^2$ and its intensity φ_0^i . The optimisation of the cost function

$$\mathcal{F}_1(\boldsymbol{\rho}, \mathbf{s}_0^1, \dots, \mathbf{s}_0^N) = \sum_{i, \mathbf{c}, \mathbf{x}} H_\varepsilon \left(I_c^i(\mathbf{p}(\mathbf{x})) - \rho_c(\mathbf{x}) \{ \mathbf{n}(\mathbf{x}) \cdot \mathbf{s}_0^i \}_+ \right) \quad (3.2)$$

is carried out by the LBFGS quasi-Newton method. The optimal $\mathbf{s}_0^1, \dots, \mathbf{s}_0^N$ are decomposed as $(\varphi_0^1, \mathbf{t}^1), \dots, (\varphi_0^N, \mathbf{t}^N)$, and $\boldsymbol{\rho}_1$ serves as a first albedo estimation.

Step 2: coarse punctual estimation. In this step, we move from far-field to isotropic near-field by “moving” the light sources from infinity to a sphere of unknown radius d . We assume that all the light sources point approximately to the centre $\mathbf{c}_S$ of the scene, and we will estimate d such that it is consistent with the previous intensity estimates φ_0^i . That is, each source will be represented by an isotropic punctual one with position $\mathbf{c}_S + d \mathbf{t}^i$ and intensity $\varphi_0^i d^2$ (consistently with the inverse-square law (2.3)):

$$\mathbf{s}_1^i(\mathbf{x}) = \varphi_0^i d^2 \frac{\mathbf{c}_S + d \mathbf{t}^i - \mathbf{x}}{\|\mathbf{c}_S + d \mathbf{t}^i - \mathbf{x}\|^3}. \quad (3.3)$$

The albedos and the distance are optimized by minimizing

$$\mathcal{F}_2(\boldsymbol{\rho}, d) = \sum_{i, \mathbf{c}, \mathbf{x}} H_\varepsilon \left(I_c^i(\mathbf{p}(\mathbf{x})) - \rho_c(\mathbf{x}) \{ \mathbf{n}(\mathbf{x}) \cdot \mathbf{s}_1^i(\mathbf{x}) \}_+ \right) \quad (3.4)$$

starting from $\boldsymbol{\rho}_1$ and an (empirical) d_1 estimate, here too using LBFGS. The optimisation result provides a new estimate $\boldsymbol{\rho}_2$ of the albedos, and some nonlinear average distance d_2 of our light sources to the centre of the scene.

Step 3: point source model refinement. We now have initial estimates of the point source positions as $\mathbf{q}_0^i = \mathbf{c}_S + d_2 \mathbf{t}^i$, intensity estimates $\varphi_2^i = \varphi_0^i d_2^2$, and a refined estimate $\boldsymbol{\rho}_2$ of the albedos. Next we “free” the positions and intensities, i.e. we use the full point source model (2.3) and optimize

$$\mathcal{F}_3(\boldsymbol{\rho}, \varphi^1, \mathbf{q}^1, \dots, \varphi^N, \mathbf{q}^N) = \sum_{i, \mathbf{c}, \mathbf{x}} H_\varepsilon \left(I_c^i(\mathbf{p}(\mathbf{x})) - \rho_c(\mathbf{x}) \{ \mathbf{n}(\mathbf{x}) \cdot \mathbf{s}^i(\mathbf{x}) \}_+ \right) \quad (3.5)$$

using LBFGS, to obtain refined values $\boldsymbol{\rho}_3$ and $\mathbf{q}_3^i, \varphi_3^i, i = 1 \dots N$.

Step 4: estimation of the LED parameters. LED anisotropies eventually come into play in sufficiently near-field situations. In that case, we can use the full LED model (2.4) to recover the full set of parameters $\boldsymbol{\theta} = (\theta^1, \dots, \theta^N)$ and $\boldsymbol{\mu} = (\mu_R, \mu_G, \mu_B)$ mentioned at the beginning of this section, optimizing:

$$\mathcal{F}_4(\boldsymbol{\rho}, \boldsymbol{\theta}, \boldsymbol{\mu}) = \sum_{i, \mathbf{c}, \mathbf{x}} H_\varepsilon \left(I_c^i(\mathbf{p}) - \rho_c(\mathbf{x}) \{ \mathbf{n}(\mathbf{x}) \cdot \mathbf{s}^i(\mathbf{x}, \theta^i, \mu_c) \}_+ \right). \quad (3.6)$$

The albedos are initialized using ρ_3 . The per-LED parameters θ^i are initialized using the previously estimated positions $\mathbf{q}_3^i$ and intensities φ_3^i , and rough estimates $\mathbf{d}_3^i = \frac{\mathbf{q}_3^i - \mathbf{c}_S}{\|\mathbf{q}_3^i - \mathbf{c}_S\|}$ of their principal directions. The anisotropy parameters are initialized using a small nonzero value $\mu_{c,0} = \delta > 0$ (starting with a nonzero value provides a better starting gradient). From the technical side, all the optimisation problems arising in the different steps of our method are solved on a GPU using JAX and OPTAX libraries, leveraging automatic differentiation for efficient gradient-based optimisation. While images of high resolution could be handled with stochastic approaches on pixels [3], we empirically observed that down-sampling images during light estimation has a negligible impact on results quality, hence we choose this alternative for very large images.

This computational framework ensures scalability and precision in solving the complex inverse problems associated with photometric and geometric reconstruction. Figure 2 illustrates the evolution of the loss function across these four different steps. As can be seen, each illumination refinement yields a better expressivity of the model, characterized by a drop in the loss function.

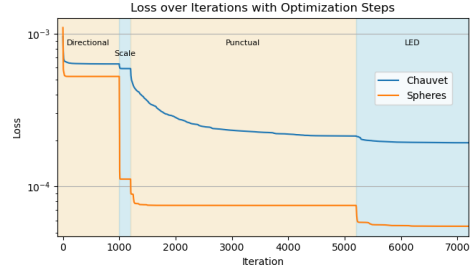


Fig. 2: Loss evolution across the four different steps, from the coarsest to the finest illumination model, on the “Spheres” and “Chauvet” datasets presented in Section 4.

Step 5: calibrated PS. Once all the light source parameters are computed, we are back in a calibrated PS setup, where we must invert Model (3.1) in terms of albedo and geometry (represented by $\mathbf{x}$ and $\mathbf{n}(\mathbf{x})$), knowing θ and μ . For this task, we can for instance keep the same objective function as in (3.6), or resort to more evolved algorithms such as variational [14] or deep [9] solvers. Let us remark that, although for LED calibration we explicitly relied on the Lambertian assumption (at least up to sparse deviations, thanks to Huber’s loss), employing deep learning-based solutions would allow us to consider much more reflective surfaces. In practice, one could thus select a subset D of “Lambertian-like” points for LED calibration, before switching to the whole set of visible points and fancy algorithms for the 3D reconstruction itself. In the preliminary experiments depicted in the next section, we however stuck to diffuse surfaces and the proposed inverse problem approach.

4 Results and Discussion

Experimental Setup. Our experimental setup comprises three real-world case studies. The images are captured using an RTI dome (reflectance transformation imaging) equipped with 105 LEDs placed on a hemisphere around the scene center. The first case involves a scene consisting of five spheres, see Figure 3. The scene geometry is perfectly known from projective geometry principles. In this case, we can observe that the light source environment is perfectly reconstructed.

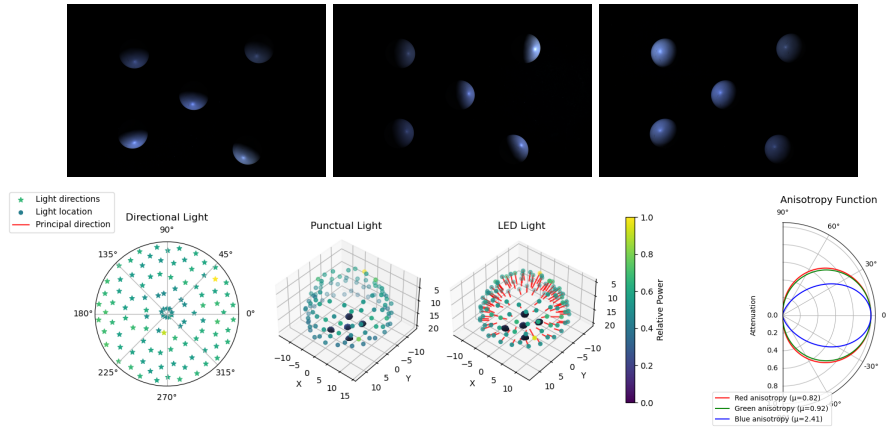


Fig. 3: Top: three (out of $N = 105$) pictures of the first scene. Bottom, left: estimated light source parameters, from the coarsest to the finest model. Bottom, right: estimated spectral anisotropy parameters.

The second case focuses on a segment of the “Panneau des Chevaux” extracted from the wall of the Chauvet cave. This segment occupies the entire image plane and is characterized by a relatively planar surface. To obtain our geometric proxy, we used the Meshroom software, which is based on SfM and MVS. The reconstruction of the illumination environment, shown in Fig. 1, is consistent with the previous one. Lastly, we consider a decimetric Cypriot ceramic head scanned under the same conditions.

Discussion on the Lighting Estimation Results. In typical scenarios, spheres exhibit a wide range of surface normals, which facilitates robust estimation at all steps of the light sources optimisation process. However, in the case of the Chauvet dataset, the curvature of the scene along the x -axis is insufficient, resulting in a predominance of coplanar normals. This lack of variation hinders the accurate estimation of directional lighting. Nonetheless, the strongly nonlinear illumination patterns (inverse square fall-off and anisotropy) produced

by LEDs on flat surfaces provide sufficient information to accurately determine both the position and the anisotropy of the LEDs. This is a particularly advantageous property, as cave walls are often characterized by planar geometry.

Additionally, as observed in the acquired images, the LEDs do not emit with uniform chromaticity in all directions. We note that it is possible to recover this directional chromatic anisotropy (bottom right image in Figure 3). This capability is especially relevant for applications involving colour-sensitive analysis, as it enables the characterisation and correction of non-uniform illumination effects.

These observations underscore the feasibility of leveraging LED-based illumination systems for reliable parameters estimation in environments with minimal geometric complexity, such as caves. They also highlight the importance of accounting for directional chromatic anisotropy in such systems to ensure accurate photometric measurements. Nonetheless, we observed in some cases, specifically with distant lights or small objects, an overfitting to the model resulting in degenerate light parameters (for example, random principal direction with null anisotropy factor). Although this can be problematic if the main interest lies in estimating the properties of the light source itself, this does not decrease the 3D reconstruction quality, as light is always evaluated on the surface of interest, where the lighting condition is optimized by construction.

Discussion on the 3D Reconstruction Results. As can be seen by comparing our results with MVS in Figure 1, the proposed method offers a markedly improved separation of colour and texture information, addressing a key challenge in the accurate reconstruction of surface reflectance properties. In particular, our technique reveals the very fine engravings surrounding these painted horses of the Chauvet cave. Furthermore, in terms of spatial resolution, the method significantly outperforms MVS, which is typically limited by its reliance on correspondences between multiple viewpoints, while the proposed approach leverages pixel-wise normal estimation of single view reconstruction.

In addition, by relying on an explicit physics-based illumination model, our approach largely reduces low-frequency bias, in comparison with alternative approaches using an implicit model. This is illustrated in Figure 4, where we qualitatively and quantitatively compare our 3D reconstruction with that obtained using a state-of-the-art Transformers-based approach [8], on the “Panneau des Chevaux” dataset. Therein, we used a 3D scanner as ground truth to evaluate the reconstructions. Although this mode of acquisition does not capture fine surface details, it provides highly reliable low-frequency geometry of the object. Consequently, a relatively high mean angular error (MAE) may partially be indicative of missing fine details in the ground truth data, in addition to the inaccuracies in the method being tested. Still, low-frequency errors are clearly visible in the state-of-the-art learning-based method.

This is further confirmed in the experiment on the Cypriot ceramic head, displayed in Figure 5. Again, low-frequency errors are much reduced, in comparison with the state-of-the-art method.

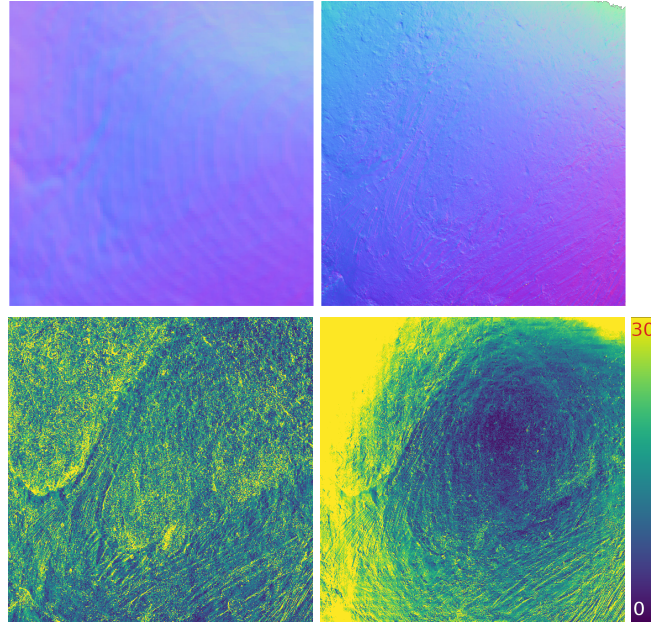


Fig. 4: Top: “ground truth” normals from the 3D scanner, and normals obtained by SDM [8]. Bottom: angular error map for our method (mean: 15.3°) and SDM (mean: 17.8°). Low-frequencies are much better reconstructed with our proposed method.

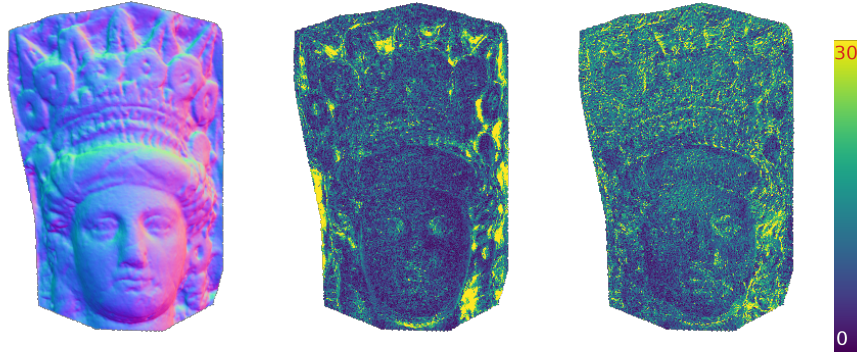


Fig. 5: Quantitative evaluation on the Cypriot ceramic head. From left to right: ground truth scanner normals, angular errors with our method (average 9.9°), and SDM angular errors (11.3°).

5 Conclusion and Perspectives

In this work, we have presented a novel uncalibrated near-light photometric stereo method, where the parameters of a physics-based illumination model for LEDs are estimated on-the-fly from a geometric proxy obtained, e.g., by MVS. A graduated optimisation procedure has been developed, which progressively refines the illumination model in a scale-aware manner. Experiments on three real datasets have shown that by combining good chromatic separation, low-frequency accuracy, and superior resolution, the method represents a notable advancement over existing techniques, offering a comprehensive solution for high-quality 3D digitisation.

On the other hand, our procedure for calibrating the illumination parameters still relies on the Lambertian model, despite the location of specularities is directly correlated with the sources parameters. In future work, we plan to further refine our approach in this direction. Moreover, the improvements obtained by the method suggest to use it in a bootstrapping process. Finally, a theoretical study of the convergence properties of our approach remains to be carried out.

Acknowledgements – We are grateful to the teams of the Chauvet cave and of the Saint-Raymond museum for allowing us to work with objects and artworks from their collections and sites, greatly contributing to this study.

References

1. Bae, G., Davison, A.J.: Rethinking Inductive Biases for Surface Normal Estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2024)
2. Brahimi, M., Haefner, B., Ye, Z., Goldluecke, B., Cremers, D.: Sparse Views Near Light: A Practical Paradigm for Uncalibrated Point-light Photometric Stereo. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2024)
3. Couprie, B., Brument, B., Laurent, A., Mélou, J., Quéau, Y., Durou, J.D.: Assessing the Quality of 3D Reconstruction in the Absence of Ground Truth: Application to a Multimodal Archaeological Dataset. In: IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (2025)
4. Favaro, P., Papadhimetri, T.: A closed-form solution to uncalibrated photometric stereo via diffuse maxima. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2012)
5. Griwodz, C., Gasparini, S., Calvet, L., Gurdjos, P., Castan, F., Maujean, B., De Lillo, G., Lanthony, Y.: AliceVision Meshroom: An open-source 3D reconstruction pipeline. In: Proceedings of the 12th ACM Multimedia Systems Conference (2021)
6. Haefner, B., Ye, Z., Gao, M., Wu, T., Quéau, Y., Cremers, D.: Variational uncalibrated photometric stereo under general lighting. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (2019)
7. Hayakawa, H.: Photometric stereo under a light source with arbitrary motion. *Journal of the Optical Society of America A* 11(11), 3079–3089 (1994)

8. Ikehata, S.: Scalable, detailed and mask-free universal photometric stereo. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2023)
9. Logothetis, F., Mecca, R., Budvytis, I., Cipolla, R.: A cnn based approach for the point-light photometric stereo problem. *International Journal of Computer Vision* 131(1), 101–120 (2023)
10. Logothetis, F., Mecca, R., Cipolla, R.: Semi-calibrated near field photometric stereo. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2017)
11. Mecca, R., Wetzler, A., Bruckstein, A.M., Kimmel, R.: Near field photometric stereo with point light sources. *SIAM Journal on Imaging Sciences* 7(4), 2732–2770 (2014)
12. Mo, Z., Shi, B., Lu, F., Yeung, S.K., Matsushita, Y.: Uncalibrated photometric stereo under natural illumination. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2018)
13. Papadhimetri, T., Favaro, P.: Uncalibrated Near-Light Photometric Stereo. In: Proceedings of the British Machine Vision Conference (2014)
14. Quéau, Y., Durix, B., Wu, T., Cremers, D., Lauze, F., Durou, J.D.: LED-based photometric stereo: Modeling, calibration and numerical solution. *Journal of Mathematical Imaging and Vision* 60, 313–340 (2018)
15. Quéau, Y., Lauze, F., Durou, J.D.: A L1-TV Algorithm for Robust Perspective Photometric Stereo with Spatially-Varying Lightings. In: Proceedings of the International Conference on Scale Space and Variational Methods in Computer Vision (2015)
16. Quéau, Y., Wu, T., Cremers, D.: Semi-calibrated near-light photometric stereo. In: Proceedings of the International Conference on Scale Space and Variational Methods in Computer Vision (2017)
17. Schönberger, J., Frahm, J.M.: Structure-from-Motion Revisited. In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (2016)
18. Schönberger, J., Zheng, E., Frahm, J.M., Pollefeys, M.: Pixelwise View Selection for Unstructured Multi-View Stereo. In: Proceedings of the 14th European Conference on Computer Vision (2016)
19. Wang, R., Yu, J., Zhang, X., Tang, Q., Liu, X.: A calibration method for LED point light sources in near-field photometric stereo. In: Proceedings of the International Conference on Optical and Photonic Engineering (2024)
20. Xie, L., Song, Z., Jiao, G., Huang, X., Jia, K.: A practical means for calibrating an LED-based photometric stereo system. *Optics and Lasers in Engineering* 64, 42–50 (2015)
21. Yang, L., Kang, B., Huang, Z., Xu, X., Feng, J., Zhao, H.: Depth anything: Unleashing the power of large-scale unlabeled data. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2024)
22. Yuille, A.L., Snow, D., Epstein, R., Belhumeur, P.N.: Determining generative models of objects under varying illumination: Shape and albedo from multiple images using SVD and integrability. *International Journal of Computer Vision* 35(3), 203–222 (1999)