A Practical and Efficient Model for Intensity Calibration of Multi-Light Image Collections

Ruggero Pintus · Alberto Jaspe Villanueva · Antonio Zorcolo · Markus Hadwiger · Enrico Gobbetti

Abstract We present a novel practical and efficient mathematical formulation for light intensity calibration of Multi Light Image Collections (MLICs). Inspired by existing and orthogonal calibration methods, we design a hybrid solution that leverages their strengths while overcoming most of their weaknesses. We combine the rationale of approaches based on fixed analytical models with the interpolation scheme of image domain methods. This allows us to minimize the final residual error in light intensity estimation, without imposing an overly constraining illuminant type. Unlike previous approaches, the proposed calibration strategy proved to be simpler, more efficient and versatile, and extremely adaptable in different setup scenarios. We conduct an extensive analysis and validation of our new light model compared to several state-of-the-art techniques, and we show how the proposed solution provides a more reliable outcomes in terms of accuracy and precision, and a more stable calibration across different light positions/orientations, and with a more general light form factor.

Keywords Multi Light Image Collections · Calibration · Light Intensity · Shape/Material Modeling · Cultural Heritage

R. Pintus, A. Zorcolo, and E. Gobbetti
Visual and Data-intensive Computing, CRS4, Cagliari, Italy
E-mail: {ruggero.pintus,antonio.zorcolo,enrico.gobbetti}@crs4.it
Web: www.crs4.it/vic

A. Jaspe Villanueva and Markus Hadwiger
Visual Computing, KAUST, Saudi Arabia
E-mail: {alberto.jaspe,markus.hadwiger}@kaust.edu.sa
Web: vccvisualization.org

Corresponding author: R. Pintus

1 Introduction

Multi-Light Image Collections (MLICs) are groups of photographs taken from the same viewpoint while changing lighting conditions. The acquired data consists of an image stack where each pixel is associated with a series of reflectance measurements. These samples are used to extract visual surface properties useful for a wide variety of tasks (e.g., exploration, data classification, surface relighting, non-photorealistic visualization), and employed in many different applications (e.g., cultural heritage (CH), natural science, industry, medical imaging) [28].

The core of all MLIC-based algorithms resides in the translation from measured reflectance changes into a set of parameters that digitally represent surface properties. Since the nature of those changes is strongly related to the light variation, a key question is if and to what extent light conditions have to be known beforehand. Although some methods aim only at a qualitative object reconstruction, to provide quantitative, reliable and repeatable outcomes, geometric and radiometric light calibration is mandatory.

Light calibration assigns an incident light direction and intensity to each MLIC measurement. Each method exploits a specific strategy, by relying on some calibration targets, and adopting a particular light model. While light positions and directions can be accurately computed by available techniques, conversely, a general, practical and reliable light intensity calibration is still an open problem. A small number of measurements on unobtrusive targets in the scene are enough to obtain a precise light position and direction. However, with an unknown light form factor, an accurate light intensity computation requires the sampling of a high portion of the camera field of view, and it typically demands for an additional capture devoted only to calibration

(e.g., flat fielding based approaches) (see Sec. 2). Unfortunately, this scenario is not applicable to some widely used capture settings (e.g., free-form MLIC acquisition).

This paper presents a novel practical and efficient light intensity calibration. We propose a hybrid mathematical solution for a new light intensity model that combines an analytical light (similarly to the standard rationale behind all physically-based calibration methods), and an interpolation-based term, which is used by image-domain calibration techniques that model the light without explicitly computing its physical parameters. In this way, we take inspiration from the strengths of those orthogonal existing methods (physically-based vs image-domain), while overcoming most of their weaknesses. The resulting approach has the advantage of both relaxing the constraint of a fixed form factor, and better minimizing the final residual error. To sum up, our two main contributions are:

- A new light model which is both general enough to be employed in a wide range of MLIC capture scenarios with different kind of illuminants, and capable of providing, at the same time, a reliable light intensity approximation with low calibration errors;
- An extensive study to evaluate the performance of the proposed solution together with several commonly used state-of-the-art approaches for MLIC-based light intensity calibration.

Synthetic and real-world experiments prove how the proposed method surpasses the calibration performances of the state-of-the-art approaches in a more general scenario.

2 Related Work

Light calibration is a wide well-known topic in the vision community. It plays an important role for a huge amount of methods/applications [28], e.g., Photometric Stereo [34], RTI [8], or SV-BRDF modeling [19]. Different illumination conditions might be considered, e.g., collinear [4], point [22], extended area lights [11], or global illumination [17]. It is out of the scope of this paper to provide an exhaustive survey of all those techniques. Conversely, we focus on the approaches strictly related to ours. The main purpose is to efficiently calibrate the geometry and the radiometric behaviour of a near light source. This is a common scenario that provides for each MLIC measurement a proper value for the light direction and intensity.

Calibrated vs Uncalibrated Methods. The first dilemma is whether to have light parameters given a-priori (e.g., obtained by a pre-calibration), or to employ an auto-calibration strategy, which simultaneously performs a specific task and estimates the incident lighting. While calibrated methods are more controlled and result in a more accurate output, uncalibrated solutions do not require any calibration target in the scene, enabling their use in the "wild". Uncalibrated Photometric Stereo [5] is a well studied research area among this type of auto-calibrated MLIC processing. To cope with the absence of a-priori knowledge of light parameters, they rely on different kind of assumptions. Some of them consider a surface with Lambertian reflectance [25,2], while others can deal with general isotropic materials [20,31], but they require an evenly spaced light constellation. Another method [18] relies on specific light positioning constraints (e.g, light along a line or in a plane), but it can only extract partial information about the scene (depth cues only). Another class of methods estimates the light direction and intensity through learning-based algorithms [6,7]. These methods exhibit a poor modeling of the light field, but they are pretty robust in the specific task of normal map estimation. However, they fail if the task is more general (e.g., appearance modeling), where both a good normal and a high quality light field are required.

Some methods try to model complex materials under unknown lighting. Dong et al. [10] recover spatially varying surface reflectance in unknown natural illumination; unfortunately, they rely on a non-fixed view acquisition (appearance-from-motion), and a-priori knowledge of object geometry. Another limitation is that they can retrieve only a single light direction/intensity per image (far point/collinear light). Papadhimitri and Favaro [26] try to solve uncalibrated Photometric Stereo with a near-light, but their light position estimation is not accurate enough (error in the order of centimeters with a 60cm light distance) to be used for further computation (e.g., SV-BRDF modeling). Huang et al. [15] propose an uncalibrated alternating minimization approach to simultaneously compute the normal map and the near-light parameters; unfortunately, they still need calibration targets in the scene, and they impose a point light model, which is not a general and common scenario (e.g., spot light). Migita et al. [23] propose a targetless optimization approach to compute shape and isotropic material properties. However, they impose a point light model without considering the distance decay factor, with an high average light direction error of 20 degrees.

For all those reasons, light direction and intensity calibration remains a mandatory step for methods and applications devoted to a high quality, quantitative shape and material modeling. There are two main approaches to light calibration in MLIC, i.e., the *Image domain* and the *Physically-based* calibration.

Image domain calibration. These methods use calibration targets to measure light direction and intensity for a small amount of pixels, and then they compute the light properties for the whole image domain through interpolation techniques. Ciortan et al. [9] extract the light directions from four pixels by using four glossy spheres, and linearly interpolate those directions across the image domain. They sample light intensity values in a white Lambertian target, and interpolate them by a cubic polynomial. Giachetti et al. [13] use the same rationale, while changing the interpolation function to a quadratic polynomial. While simple and practical in many scenarios, these methods lack accuracy, since they completely ignore the geometry of the light (position or orientation), and the radiometric behaviour of the light field (distance or angular attenuation). They can manage non-collinear light, but the closer the illumination becomes (near-light field), the less the 2D interpolation is reliable. The resulting error strongly impact the MLIC processing such as normal computation, BRDF fitting, relightable image modeling. Other methods [32,3] aim at finding a data driven light vector field. They employ a flat fielding correction obtained by the acquisition of a calibration target before the actual capture. While extremely accurate, this is not generally applicable to all types of MLIC acquisitions (e.g., a free form case, when it is not possible to do two equal captures); moreover, even with fixed light domes, it does not consider intensity repeatability issues between one acquisition and the other, which is not rare with low-cost illuminants.

Physically-based calibration. In this class of methods an analytical light model is imposed, and the information measured through the calibration targets are used to find its parameters. Analytical light modeling is a well studied field, with seminal approaches that represent the image formation model of linear illuminants [16]. Others couple the point light model with a perspective camera to solve near-field Photometric Stereo [22]. Similarly, Huang et al. [15] employ a point light with the distance fall-off and camera vignetting. Xie et al. [35] and Quéau et al. [29] present a LED-based calibration that includes a decay of light intensity driven by the angle with the optical axis. They assume a fixed LED light, whose parameters are known a-priori, and only its position, direction and intensity have been calibrated. Pintus et al. [27] address this limitation with a non-linear optimization to find the unknown optical axis and the exponential decay parameters. Similarly, Ma et al. [21] use a perspective camera and a Lambertian calibration plane to estimate the non-isotropic Radiant Intensity Distribution (RID) of a near point light source. While these methods are very accurate, most of them are tailored to a specific light form factor, and they

are not generally applicable. Some are computationally expensive with slow non-linear optimizations. To ensure convergence, some have to capture many samples across the image, which requires a separate acquisition only for the calibration; again this makes those methods hardly applicable in a MLIC free-form scenario.

Our contribution. The proposed technique takes inspiration from those extreme strategies. Our hybrid approach takes the best characteristics from both, and results in a more practical, efficient and generally applicable method. We take into account physically-based light features (3D position, direction, intensity distance decay) while keeping the light model extremely simple (dimensionless point light). Conversely, we take insights from the image domain methods. Rather than blindly interpolating the measures on the calibration target, we first remove all the physically-based contributions modeled by our simple light source, and then we show how the residual error is a more tractable data that can be efficiently represented by image domain interpolation. Our light calibration can be efficiently solved by linear optimization, and, being numerically more stable, does not require a dense sampling of the light field. This results in a smaller calibration target that can be easily inserted in the scene without decoupling the calibration and the actual MLIC acquisition; this avoids repeatability issues across different captures. Moreover, it can be generally applicable to different MLIC setups (from fixed domes to free-form) and light types.

3 Method

We present here a general and practical model for calibrating the spatially varying intensity of a light source. By performing some measurements on a diffuse target of known reflectance, the model is capable of assigning the proper light intensity value to each pixel in each MLIC image.

First of all, we consider a Lambertian diffuse target, which has an optical response equal to:

$$w(u,v) = \frac{1}{\pi} \rho(u,v) \, \phi(u,v) \, \frac{x_s - x(u,v)}{\|x_s - x(u,v)\|} \cdot \hat{\mathbf{n}} \qquad (1)$$

where w is the measured value within the white target at the image pixel coordinates (u,v), ρ is the target response, ϕ is the light intensity, x_s is the $3D$ position of the light, x is the $3D$ position of the measured point on the white target corresponding to the pixel (u,v), and $\hat{\mathbf{n}}$ is the planar normal. From the light position calibration step x_s, x and $\hat{\mathbf{n}}$ are known. We use a common white paper with a known constant ρ.

Unlike previous approaches, we propose to model the term ϕ as the combination of both a physically-based

term and a $2D$ interpolation function (for simplicity, when it is not needed, we drop out the u, v dependency):

$$\phi = \phi_I \phi_P \qquad (2)$$

The physically-based term ϕ_P is modeled as a simple isotropic, point light source, and it takes into account the physical position of the light in $3D$, its absolute light intensity, and the light intensity decay due to the squared distance between the light and the illuminated point in space. We can write this term as:

$$\phi_P = \frac{\phi_0}{\|x_s - x\|^2} \qquad (3)$$

where ϕ_0 is the absolute light intensity at a unitary distance from the light. The general spatially varying nature of the light (e.g., spot like model) is captured by the term ϕ_I, which we model within the image domain. Depending on the type of interpolation function, this term is explicitly a function of the pixel (u, v) or of the light direction $\hat{\mathbf{d}} = x_s - x / \|x_s - x\|$ (which in turn depends on the image pixel coordinate in an indirect way). Moreover, different formulations exhibits a different set of parameters $\mathbf{p}$. So, the final light intensity model will be:

$$\phi = \frac{\phi_0}{\|x_s - x\|^2} \phi_I (\cdots, \mathbf{p}) \qquad (4)$$

where $\cdots$ might be the $2D$ image domain (u, v) or the $2D$ hemispherical domain of light directions $\hat{\mathbf{d}}$. Beside the specific parameterization, our idea is to define ϕ_I as a general term expressed with a linear combination of basis functions. We can generalize the mathematical representation of ϕ_I as:

$$\phi_I (\cdots, \mathbf{p}) = \sum_i p_i \beta_i (\cdots) \qquad (5)$$

where β_i are the basis functions. This formulation has the advantage of being very efficient, since it involves only a linear system to calibrate the light. Moreover, it can be generalized to a wide range of light types depending on the capability of the functions β_i to represent a general spatially varying light field. To the extreme, the term ϕ_I might be not only a set of more complex bases or functionals, but also it can be expressed as a generalized mapping between a 2D field (image pixels or light directions) into a scalar field (light intensity), e.g. through a neural network implementation.

For the sake of presentation and its clarity, among all the possible implementations in this paper we concentrate on five possible formulations of the ϕ_I term, ranging from the most simple, straightforward solutions (e.g., linear interpolation), to more complex bases. We call these solutions with the prefix *Residual-*, since one

possible interpretation of this term is that it interpolates, in a $2D$ manifold domain, a residual error done when trying to fit a physically-based model. In particular, we explore here two main approaches. The first is a polynomial based interpolation; we consider three implementations of it, i.e., a linear interpolation (*ResidualLinear*), a quadratic term (*ResidualQuadratic*), and a cubic function (*ResidualCubic*). The form of those terms respectively is:

$$\phi_I (u, v, \mathbf{p}) = p_0 u + p_1 v + p_2 \qquad (6)$$

$$\phi_I (u, v, \mathbf{p}) = p_0 u^2 + p_1 v^2 + p_2 uv + p_3 u + p_4 v + p_5 \qquad (7)$$

$$\phi_I (u, v, \mathbf{p}) = p_0 u^3 + p_1 v^3 + p_2 u^2 v + p_3 uv^2 + p_4 u^2 + \\ + p_5 v^2 + p_6 uv + p_7 u + p_8 v + p_9 \qquad (8)$$

The second approach is derived by the consideration that, once we fix the light position and we remove the quadratic dependency on the distance (by using the physically based term ϕ_p), we can see that interpolating across the domain (u, v) is equivalent to interpolate in the light direction domain across the hemisphere above the object. For this reason, a good choice for ϕ_I is a set of bases specifically designed to approximate functions defined on a sphere or a hemisphere. We propose here two alternatives, i.e., the *Spherical Harmonics* basis, and the so called *hemispherical basis* or *H-Basis* [14]. While keeping the basic advantages of *Spherical Harmonics* and requiring the same amount of coefficients, the *H-Basis* is designed to represent irradiance signals over the hemisphere of possible surface normals, and generally exhibits less error than other hemispherical bases. The two light models we propose that are derived from these two bases are the *ResidualRSH* (RSH stands for Real Spherical Harmonics), and the *ResidualHBasis* approaches, and the resulting ϕ_I are:

$$\phi_I(\hat{\mathbf{d}}, \mathbf{p}) = \sum_{l=0}^{N} \sum_{m=-l}^{l} p_l^m Y_l^m(\hat{\mathbf{d}}) \qquad (9)$$

$$\phi_I(\hat{\mathbf{d}}, \mathbf{p}) = \sum_{i=0}^{N} p^i H^i(\hat{\mathbf{d}}) \qquad (10)$$

where $Y_l^m(\hat{\mathbf{d}})$ and $H^i(\hat{\mathbf{d}})$ respectively are the *Real Spherical Harmonics* and the *H-Basis* functions.

3.1 The calibration pipeline

After having introduced our new light model, we describe now how to use it to perform the actual light calibration. This consists in three main parts, the prerequisites and the setup, the light positions estimation, and the light intensity calibration.

Fig. 1: MLIC acquisition setup. The camera points toward the object, surrounded by some calibration targets (glossy spheres and a white frame).

3.1.1 Prerequisites

The prerequisites of our calibration method are the same as in standard MLIC pipelines. We share the same acquisition devices, scene setup, object topology, capture procedure, and type of input data. In particular, similarly to other light intensity calibration methods [9, 13, 27], our solution needs to sample the light intensity in a small portion of the image, and on a planar target of known optical response. In order to do that, we also require the information about light positions in the camera reference frame. These requirements are general enough that can be met by standard MLIC capture settings and pre-processing routines. Moreover, among the standard capture procedures, we consider the one that is the most challenging from the calibration point of view, i.e., the free-form MLIC acquisition, where the light is hand-held and freely moved around the object. This setup does not allow a fixed and more precise calibration; conversely, it requires that a light calibration is run for each MLIC image independently, and it is not repeatable from one acquisition to another.

Without loss of generality, among all the available, standard free-form MLIC setups, we choose the one with the object under study surrounded by both some glossy spheres (for light geometry calibration), and a white planar frame (for light intensity calibration); we adopt this setup since it is easy to realize and it is the most common one (see Figure 1). As usual, all the involved elements are designed to build a framework that is extremely simple and usable by non-experts with a minimal training.

After the acquisition, the input of our calibration are a series of captured raw images, the intrinsic parameters of the camera lens, and the radius of the glossy spheres.

3.1.2 Light position calibration

Before modeling the light intensity, we need to extract some geometrical information from the acquired data. For the results presented in this work, we adopt the following workflow. We undistort the original images and segment the glossy spheres. The projection of a sphere onto an image is a $2D$ conic. We compute the sphere $3D$ positions by combining camera intrinsic parameters and the equation of the $2D$ conics [33]. Since the spheres are on the planar frame, from the $3D$ conics we can compute the plane equation. After these steps, for each pixel belonging either to a sphere or to the planar target we are capable of launching a ray and finding the corresponding $3D$ position. Given the $3D$ positions of the spheres, we use the technique proposed by Ackermann et al. [1] to compute the $3D$ position x_s of each light in the MLIC.

3.1.3 Light intensity calibration

The knowledge about the camera parameters, the geometry of the scene, and the light position allow us to compute, for each pixel in the planar target, the $3D$ position of the point on the plane, the direction of the light ray, and the distance between the light source and that point. The information we miss is the intensity of the light at that point. From the equations above, we can express the single measurement k done on the planar target as:

$$w_k = \frac{1}{\pi} \rho \left(\frac{x_s - x_k}{\|x_s - x_k\|^3} \cdot \hat{\mathbf{n}} \right) \phi_0 \sum_i p_i \beta_i \left(\cdots_k \right) \qquad (11)$$

Calibrating the light intensity means solving a system of equations to find the values of ϕ_0 and p_i. Generally, given K measurements, we can consider ϕ_0 as the following average:

$$\phi_0 = \frac{1}{K} \sum_k \left[\phi_0 \sum_i p_i \beta_i \left(\cdots_k \right) \right] =$$
$$= \frac{1}{K} \sum_k \frac{w_k \pi}{\rho} \frac{1}{\frac{x_s - x_k}{\|x_s - x_k\|^3} \cdot \hat{\mathbf{n}}} \qquad (12)$$

and we can compute the light parameters p_i by solving a simple linear system $B\mathbf{p} = \mathbf{w}$, where B is the matrix with the values of the basis functions, and $\mathbf{w}$ is the known term that includes the measurements, ϕ_0, and all the geometrical terms.

4 Results

We validate the proposed solution by comparing its performances with those of various techniques that are both commonly used calibration strategies or state-of-the-art approaches. We separate this analysis into two main phases. First, we conduct a series of synthetic experiments to test the proposed model in a completely controlled manner (Sec. 4.1). Then we investigate if the synthetic results are confirmed in several kind of real-world scene acquisitions and calibrations (Sec. 4.2).

4.1 Synthetic tests

To render the synthetic MLICs, we need to define a synthetic scene, with a reference frame, some camera parameters, some virtual calibration targets, an illumination source with a series of positions/orientations (one for each MLIC image). Without loss of generality, for all tests we consider the camera centered at the origin of the reference frame, and pointing to the negative z-axis. The camera has a field of view of about 60 degrees and a resolution of 1845×1232 pixels, with a centered principal point. The virtual target is a grey Lambertian diffuse plane; its normal is collinear with the frame z-axis and it is positioned in front of the camera at a distance of about $400cm$. Across the experiments we modify the number of lights, their types, and their attributes (position, orientation, intensity). We employ 5 light types: *Isotropic*, an isotropic point source; *Area*, a spherical area light (radius equal to $1cm$); *SpotBW00*, a spot light with linear axial decay and no central beam of constant intensity (*BW* stands for Beam Width); *SpotBW10*, a spot light with linear axial decay and a 10 degree central beam of constant intensity; *LambertLED* [29], a Lambertian LED. Although synthetic, to obtain a high quality, realistic rendering of our scenes, we employ the physically-based renderer *Mitsuba* [24], and we render high-dynamic range images with half floating point precision. Figure 2 shows some MLIC images rendered with our 5 light types positioned at the same location.

We test nine calibration algorithms: *Collinear*, *Point*, *Spot* [27], *Quadratic* [13], *ResidualLinear*, *ResidualQuadratic*, *ResidualCubic*, *ResidualRSH*, and *ResidualHBasis*. *Collinear* is the most common and widely used directional approximation of light model. *Point* is the analytical model of an isotropic point light. *Spot* is a spot light with an axial decay modeled as $\phi(\theta) = \phi_0 cos^\mu(\theta)$ [29, 27], where ϕ_0 is the intensity at the light axis, θ is the angle between the emitting direction and the light axis, and μ is the exponential falloff. *Quadratic* is the image domain interpolation of light intensity by using a quadratic polynomial. The methods starting with *Residual* are

based on the proposed approach, and they respectively model the residual with a linear, quadratic, or cubic polynomial, Real Spherical Harmonics (RSH) [30], or H-Basis [14]. All the calibration techniques use only a small subset of all the pixels in the planar target (*Training* set) located at the edge of the image. In this manner, we want to simulate a real acquisition, where the number of pixels belonging to calibration targets are minimized (and are typically at the edge of the image). We employ the remaining pixels to check the calibration error (*Test* set). We perform a calibration for each single MLIC image independently. After the estimation of the light intensity parameters, for each *Test* pixel we compare the original rendered value with that predicted by the calibrated model. Then, we compute some error statistics and error maps, in order to evaluate and rank the algorithms' outcomes.

We present four different synthetic tests. In the first, the MLIC has been built with the classic structure of a dome light constellation. We place 52 lights evenly distributed across an hemisphere of $30cm$ radius. In the second experiment, we investigate calibration performances from a very near to a far illuminant condition; while keeping a chosen direction constant, we vary the distance between the light and the center of the target within a range from $20cm$ up to $2.7m$. The third test aims at showing the calibration performance as a function of the main incident angle. We build a 16 image MLIC with a fixed light distance (about $330cm$), and we change the zenith angle from 5 to 80 degrees. Finally, we take the first 52 light MLIC, and we test the robustness of the calibration techniques by simulating an error in the light position calibration. The distorted light positions are obtained by randomly moving each light from zero to one centimeter away from the correct positions.

For each test, we produce five MLICs (each for a different type of light), and we launch the nine calibration techniques on them. We then evaluate the quality of the calibration by computing the average relative error between the original image and the predicted one. The relative error is expressed as:

$$e_r = \frac{1}{N} \sum_\Omega \frac{|p - \tilde{p}|}{p} \tag{13}$$

where p and $\tilde{p}$ respectively are the ground truth and the predicted pixel value, Ω is the domain of *Test* pixels, and N is the number of pixels. For each test we then plot the cumulative average relative error across all pixels, all MLIC images, and all calibration techniques. The average relative error is expressed as a percentage in logarithmic scale, and we include some vertical lines to depict the minimum and maximum errors. We see that,

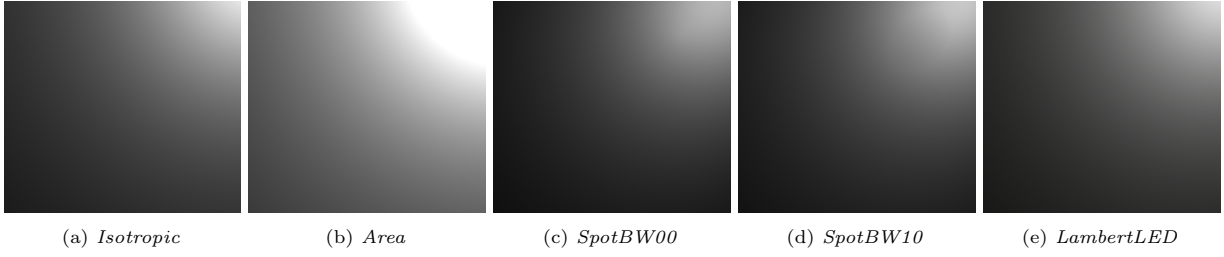(a) *Isotropic* (b) *Area* (c) *SpotBW00* (d) *SpotBW10* (e) *LambertLED*

Fig. 2: MLIC Images. Some example of MLIC images created by a physically-based rendering engine (*Mitsuba*) under different light types. The rendered object is a virtual planar target of known optical response.
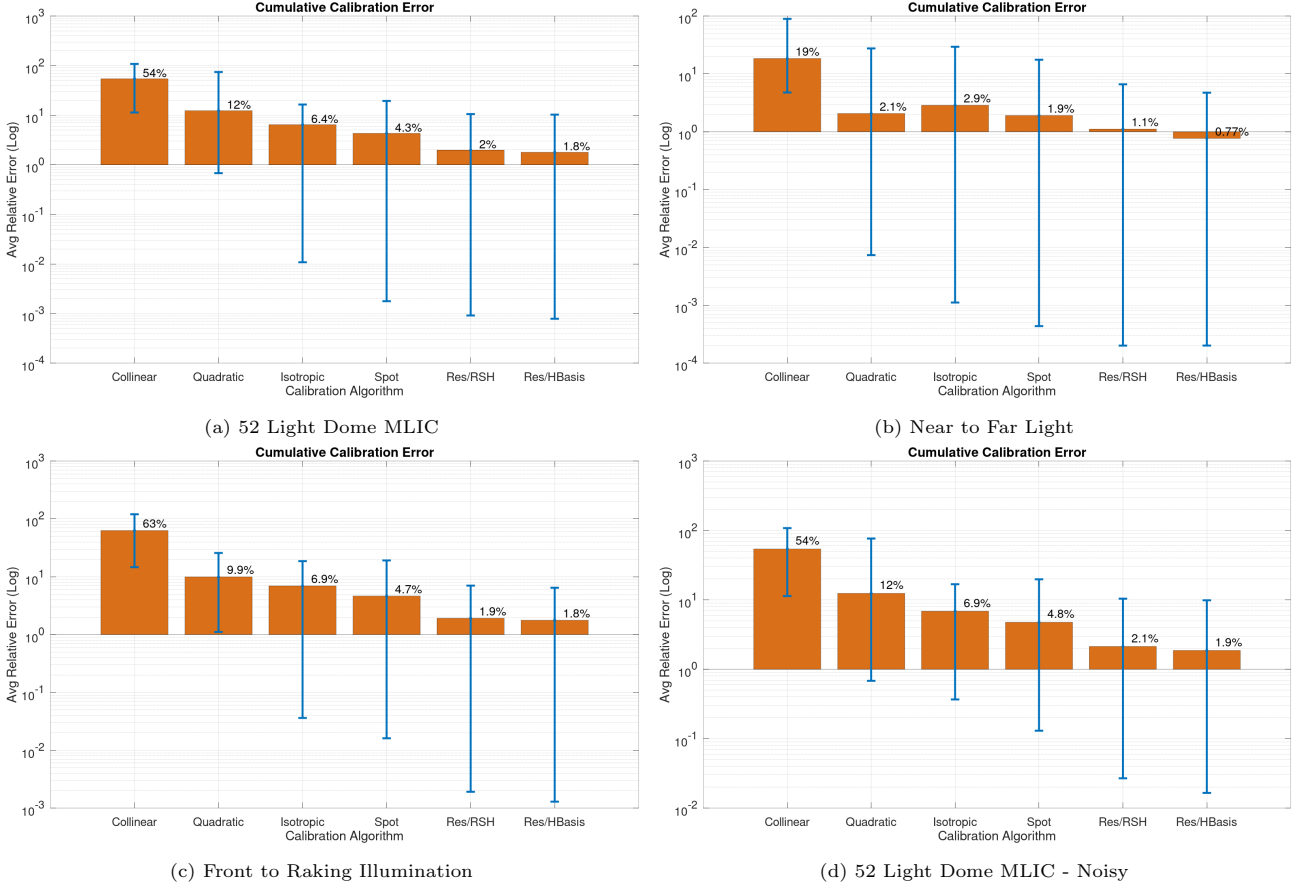

(a) 52 Light Dome MLIC

(b) Near to Far Light

(c) Front to Raking Illumination

(d) 52 Light Dome MLIC - Noisy

Fig. 3: *ResidualRSH* and *ResidualHBasis* vs state-of-the-art techniques. We compare the cumulative average relative error values of the *ResidualRSH*, the *ResidualHBasis*, and all the other techniques, computed across all the images within the MLICs of all light types. Vertical lines depict the minimum and maximum relative errors. Each plots refers to a different experiment. (a) a 52 light dome-like MLIC. (b) the error as a function of the distance between the light and the center of the target; we fix the light direction and vary the distance from $20cm$ to $2.7m$. (c) the error as a function of the zenith angle (from 5 to 80 degrees). (d) the same 52 light dome-like MLIC as before, where we randomly apply a noise/bias ranging from zero to one centimeter to all the light positions. *ResidualRSH* and *ResidualHBasis* exhibit the best performances across different capture scenarios.

among *Residual*-based algorithms, *ResidualRSH* and *ResidualHBasis* are always better than the others. So, to avoid too cluttered plots, from now on we compare only these two best ones with the other state-of-the-art methods.

Figure 3 shows the error statistics for the four experiments. Of course, *Collinear* is definitely the worst in the case of a near light scenario; in fact, we include it mostly as a "control" algorithm, a sort of upper bound for the near light calibration error. *Quadratic* [13] is in general the second worst method; this method can exhibit per-image relative errors up to 70%, because it is not capable of modeling the intensity mostly in raking light conditions. *Point* is a very simple approach, and it proved to
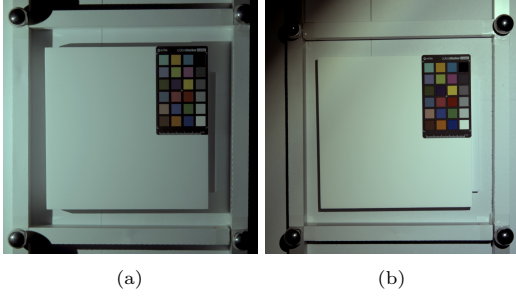
(a) (b)

Fig. 4: Spectralon and Color Checker MLIC. (a) one image from the MLIC in the first experiment (Sec. 4.2.1). In the case of a non-centered light axis (Sec. 4.2.2), we capture a MLIC by deliberately orienting the light axis off the center of the scene. On the right (b) we show how the object is still illuminated by the LED light but the light cone is not centered. This creates a less homogeneous light intensity across the image.

have a very stable behaviour; however, its error is bigger, because it fails to model all the *Spot* or more complex light types. *Spot* [27] can reach very low errors with all the types of spot light. Unfortunately, while this error is less than the previous ones, it remains high due to two main reasons: it is not capable of dealing with lights that slightly differ from the chosen analytical model; even in the best case, its calibration is very unstable, since the method proposed by Pintus et al.[27] has an uncertainty in the initialization of the light axis, which influences the resulting convergence of the optimization routine, and affects the final local minimum of the light parameters fitting. In the best cases, it produces very low errors (comparable with those of *ResidualRSH* or *ResidualHBasis*), but, as long as we move toward the front light conditions its performance are comparable to *Point*. The weaknesses of *Point* and (mostly) *Spot* are evident in the noisy case, where those methods increase by a relative error of about 0.5%. *ResidualRSH* and *ResidualHBasis* prove to be more reliable and stable across different light types and light positions, so that, at the end, they keep a very low cumulative error.

4.2 Real-world scenes

We present five tests performed in a real-world setting. In the first three experiments we capture an object of known reflectance to quantitatively evaluate the calibration approaches. In the last two tests, we measure the calibration performances in terms of accuracy and precision in estimating the surface normal and albedo.

4.2.1 Spectralon

The first real-world scene consists of a white target calibration frame, a *Spectralon* [12], and a color checker. The
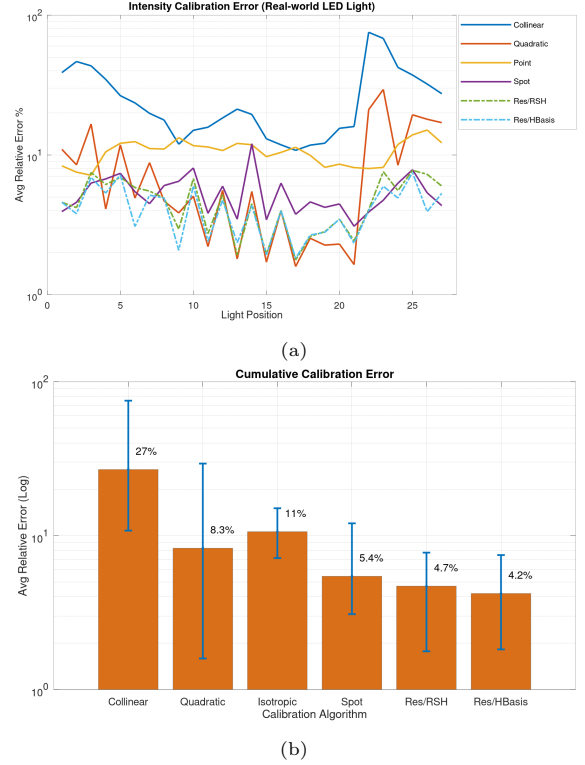

(a)

(b)

Fig. 5: Spectralon MLIC. Average relative error values computed (a) for each image, and (b) across all MLIC images and all light types. Vertical lines are minimum/maximum errors. *ResidualRSH* and *ResidualHBasis* prove to be more stable/reliable across different lighting conditions.

frame around the Spectralon is used for intensity calibration (*Training* pixels), while *Test* pixels are those on the Spectralon. The target is made of common white paper, with a response ρ of $RGB = \{0.783, 0.798, 0.835\}$; the *Spectralon* has a reflectance response of about 0.99 in all the visible spectrum. The MLIC has been acquired by freely moving a LED light in the hemisphere above this scene at a distance of half-meter. All images have been captured with a 14*bit* precision. One image of the scene is shown in Figure 4a.

We compute the average relative error for each image and each calibration strategy (see Figure 5a). We report also the cumulative relative error across all pixels and all images (see Figure 5b). *ResidualRSH* and *ResidualHBasis* exhibit the lowest average and maximum error. It is about half of *Quadratic*. Although *Spot* resembles the most the LED light used in the acquisition, *ResidualRSH* and *ResidualHBasis* are more stable across different light positions/directions. For the lights with a zenith angle bigger than 60 degrees (#22 ∼ #27), *ResidualRSH* and *ResidualHBasis* produce a quite smaller error than *Quadratic*. For those angles they are comparable with *Spot*, which conversely tends to fail for lights coming from above (#14, #9 or #15 ∼ #20).

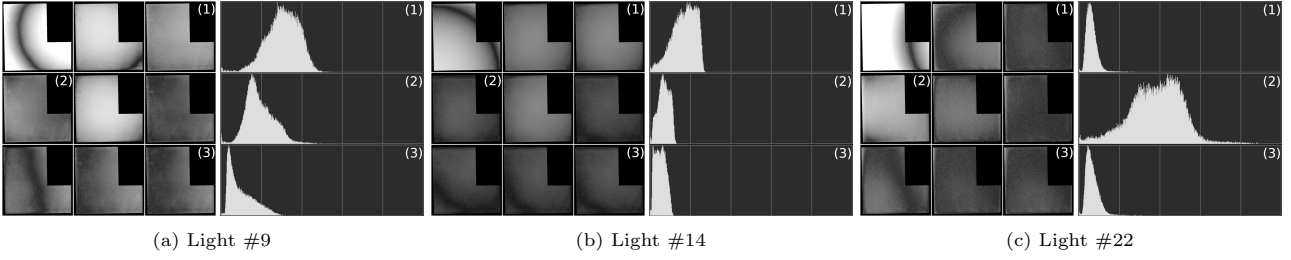(a) Light #9   (b) Light #14   (c) Light #22

Fig. 6: Relative error maps. For three exemplifying light positions, we show the relative error maps by using the 9 calibration strategies organized in a 3x3 grid (left to right and top to bottom): *Collinear*, *Point*, *Spot* (#1), *Quadratic* (#2), *ResidualLinear*, *ResidualQuadratic*, *ResidualCubic*, *ResidualRSH*, *ResidualHBasis* (#3). The error histograms for methods (#1), (#2) and (#3) prove how *ResidualHBasis* (*ResidualRSH* behaves similarly) is more stable/reliable across different lighting conditions.

This happens when convergence problems arise when a less foreshortened relationship exists between the light beam and the white target. In this case, the sampled light intensity variation is not sufficiently heterogeneous and complete for a robust *Spot* light fitting.

Figure 6 shows the spatially-varying relative error maps across the *Spectralon* pixels for three MLIC lights. We adjust the luminance levels for visual clarity sake. The 3x3 grid on the left shows the relative error maps for all the nine calibration approaches (left to right and top to bottom): *Collinear*, *Point*, *Spot* (#1), *Quadratic* (#2), *ResidualLinear*, *ResidualQuadratic*, *ResidualCubic*, *ResidualRSH*, *ResidualHBasis* (#3). We analyze the histograms of the (#1), (#2) and (#3) luminance relative error. We omit *ResidualRSH* from the histograms for clarity, since its performance is very similar to *ResidualHBasis*. Figure 6a depicts how *ResidualHBasis* increases the performance of both *Spot* and *Quadratic*. Figure 6b and Figure 6c show how *ResidualHBasis* is more stable than *Quadratic* or *Spot*, which sometimes are comparable with *ResidualHBasis*, but they can be far worse than it for some lights. This confirms the same trend in the synthetic data and in the Figure 5a, where *ResidualRSH* and *ResidualHBasis* prove to be more stable/reliable across different pixels and different lighting conditions.

We include the color checker to ensure that the intensity calibration methods ensure a correct white balance. *Spot*, *ResidualHBasis* and *ResidualRSH* exhibit similar (and good) white balance and color calibration performance.

4.2.2 Non-centered light axes

In this experiment we select only a single azimuth angle and a small set of zenith angles, and we test the performances by varying the orientation of the light axis. In particular, rather than pointing the light to the scene center, we deliberately orient it in an extreme manner (random for each light and far from the scene center), with the only constraint of the scene being within the
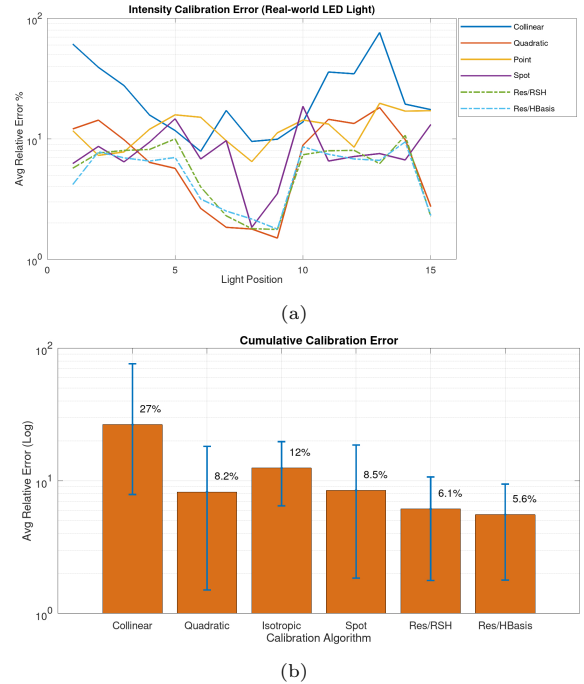

(a)

(b)

Fig. 7: Non-centered light axis. Average relative error values computed (a) for each image, and (b) across all MLIC images and all light types. Vertical lines are minimum/maximum errors. *Quadratic* and *Spot* exhibit an error up to 20%, and up to five times bigger than the proposed approach. *ResidualRSH* and *ResidualHBasis* prove to be more stable/reliable across different lighting conditions.

light cone. Figure 4b shows this setting; in the upper left part we can see the edge of the light cone, and the scene is now illuminated in a less homogeneous manner.

In the case of the most raking lights (#1 ∼ #3, #11 ∼ #13), *Quadratic* performs worse than *ResidualRSH* and *ResidualHBasis*, and exhibits a maximum error of about 20%. For high elevation angles *Spot* tends to fail more; for light #7 it is five times *ResidualRSH* or *ResidualHBasis*, and for light #10 the error is about 20%. Sometimes (#14), *Spot* is comparable with *ResidualRSH* or *ResidualHBasis* even if it is a front light,
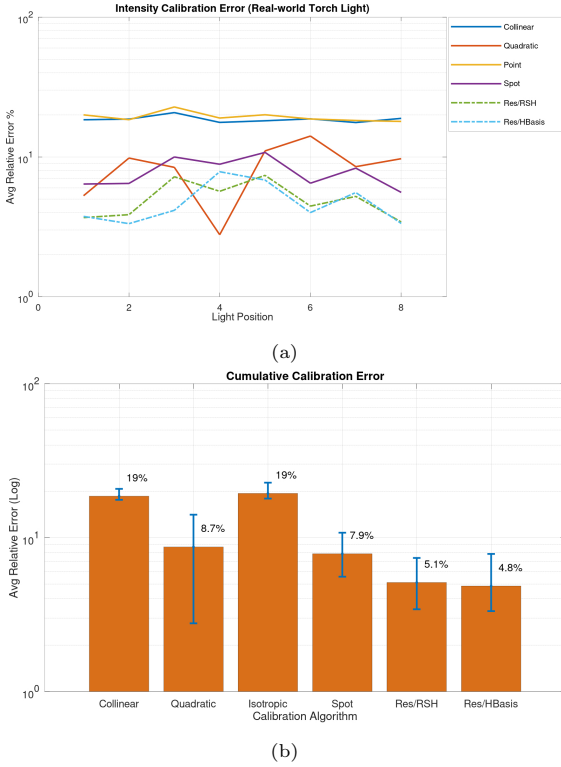
(a)

(b)

Fig. 8: Torch lamp. Its light field cannot be easily modeled by the point, spot, LED or other known types of lights. In this scenario, our approach proves to better approximate this more general light form factor.

because the high variability of light intensities in the edge of the light beam (Figure 4b) helps the *Spot* fitting algorithm in finding the right LED axis. Nonetheless, in general, we can confirm that both *ResidualRSH* or *ResidualHBasis* remain more stable/reliable while varying light position and direction.

4.2.3 Real-world torch

Another advantage of the proposed light model is that we can deal with an unknown, more general light form factor. We present a MLIC obtained with a common torch lamp, which consists of a LED light and a group of lens in front of it. This cannot be easily modeled by a point, spot, LED or other known types of lights. By looking at the standard MLIC practices (e.g., CH capture settings [8]), this is not an unusual, rare scenario. Like the previous tests, we acquire a MLIC and we apply all the calibration algorithms. Figure 8 shows both the per image average relative error and the cumulative error across all pixels and light positions. Since the considered spatially varying light intensity cannot be modeled by any of the physical models, it is evident how in this experiment *ResidualRSH* and *ResidualHBasis* methods exhibit better performances. While *Spot*, although not

Method	Albedo		Normal Map	
	RMSE Luminance	PSNR	RMSE Degree	PSNR
Collinear	0.174	21.23	16.177	17.05
Quadratic	0.058	30.82	7.981	23.16
Isotropic	0.108	25.39	3.580	30.11
Spot	0.031	36.10	2.341	33.79
Res/RSH	**0.015**	**42.51**	2.389	33.62
Res/HBasis	0.019	40.35	**2.250**	**34.14**

Table 1: Accuracy. Error between the ground truth albedo and normal maps and those maps computed with the calibration methods. We compute the error in terms of RSME of the albedo luminance and the angle deviation in degrees, and in terms of PSNR. *Residual*-based approaches exhibit the best error statistics.

stable, in previous tests had performance similar to our methods, here it is always worse than them. Moreover, the minimum error of *Spot* is bigger than the average of both *ResidualRSH* and *ResidualHBasis* (Figure 8b)

4.2.4 Accuracy (Fitting vs Flat-field)

To assess the accuracy of the several calibration approaches, we need to provide some sort of ground truth data. We produced a calibrated MLIC by employing a flat-field calibration. This approach consists in two MLIC acquisitions with exactly the same lights (we use a fixed light dome). In the first acquisition, a *Spectralon* is used to measure the actual per-pixel light intensity. Then, a second acquisition is performed by replacing the *Spectralon* with the object we want to capture (see Figure 9a). After that, we use this calibrated MLIC to compute two surface properties, i.e., the albedo and the normal map; we use these as two ground truth maps. In the setup used to acquire these MLICs we also include a white frame. We use this frame for two reasons: first, to apply all the calibration methods to the second capture only (without the information from the *Spectralon*); second, to compensate slight differences of light brightness due to non-perfect repeatability of light conditions in the two consecutive captures. For each calibration method we compute the corresponding albedo and normal map. Finally, we measure the accuracy of each method by computing the error between their albedo/normal maps and the ground truth albedo/normal map. Table 1 presents the error statistics. For the albedo, we compute the RMSE of the difference Luminance, and the PSNR. For the normal map, we compute the RMSE of the angular deviation in degrees, and the PSNR. *Residual*-based calibrations exhibit the best error performance. In the albedo, the RMSE of *ResidualRSH* is about 50% of *Spot*, and 25% of *Quadratic*. For the normal map, *Residual*-based methods exhibit an error that is one third of *Quadratic*, while comparable with *Spot*.
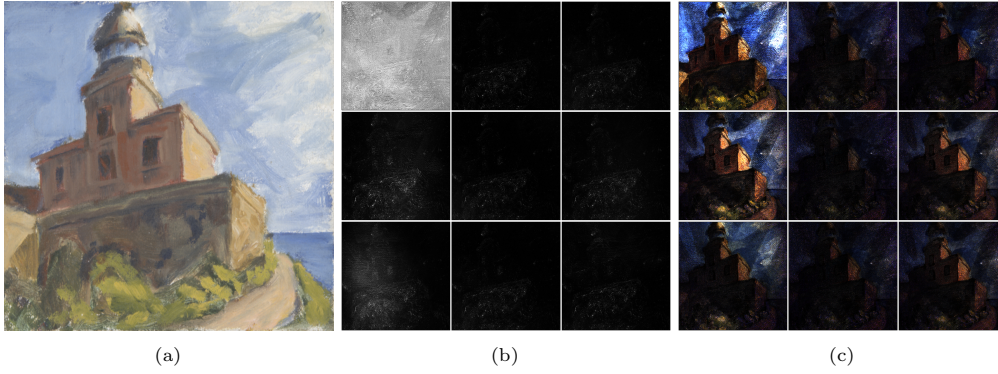
Fig. 9: Repeatability. We acquire five MLICs of the object in (a). For each calibration method and for each MLIC we compute (b) the normal map and (c) the albedo map, and then we estimate the per pixel standard deviation. The higher is the standard deviation, the less repeatable is the calibration method, the less precision it exhibits. The 9 calibration strategies are organized in a 3x3 grid (from left to right and top to bottom): *Collinear*, *Point*, *Spot*, *Quadratic*, *ResidualLinear*, *ResidualQuadratic*, *ResidualCubic*, *ResidualRSH*, *ResidualHBasis*. *ResidualRSH* exhibits the best performance in terms of precision.

Method	Albedo Standard Deviation (σ)	Normal Map Standard Deviation (σ)
Collinear	0.0124	0.0642
Quadratic	0.0103	0.0164
Isotropic	0.0069	0.0111
Spot	0.0071	0.0110
Res/RSH	**0.0062**	**0.0109**
Res/HBasis	0.0077	0.0117

Table 2: Precision. Average Standard deviation computed from five MLIC captures of the same scene. Low values of Standard deviation mean high precision.

4.2.5 Precision (Repeatability)

In the previous test we have analyzed the performance of different calibration approaches in terms of accuracy. Now we would like to test the precision (or measure repeatability). We take again the object in Figure 9a and we acquire five different MLICs. For each calibration strategy, we first calibrate the light (direction and intensity), and then we process the five MLIC to compute five surface normal and albedo maps. Now we evaluate the level of repeatability by computing the standard deviation of the five normal maps (or albedos). The precision of each calibration is proportional to the inverse of that standard deviation. Figure 9b and Figure 9c respectively show the nine standard deviations of the normal and the albedo map; for the sake of clarity, we apply a tone mapping to make the image contrast higher. The 3x3 grid depicts the methods with the same order presented in Figure 6. *Collinear* method has not only the lowest accuracy but also the lowest precision. Similarly, *Quadratic* has a high standard deviation (i.e., low precision). While *Point* exhibits a high absolute error (low accuracy) in the previous tests, its precision is among the lowest, comparable with *ResidualLinear*, *Residual-RSH*, and *Spot*. *Spot* light proved to be not so reliable

in the absolute intensity measure (low accuracy), while here its precision is the better than *ResidualHBasis*, and comparable to *ResidualRSH*. Together with the previous experiments, we can say that *ResidualRSH* and *ResidualHBasis* prove to be the most stable/reliable in terms of accuracy, while *ResidualRSH* also exhibits the best performance in terms of precision.

5 Conclusions

We have presented a novel practical and efficient method for light intensity calibration. The proposed illumination model is very simple, and results in a mathematical formulation that relies on computationally efficient linear solvers. The combination of a fixed physically-based term and an interpolation function that minimizes the residual error makes it possible not to impose a fixed light type; this makes our method very versatile, and extremely adaptable to different setup scenarios. The analysis of the performance of our method compared to the most used light intensity calibration strategies highlights how the proposed solution advances the state-of-the-art in terms of accuracy and precision of both light intensity fitting, benefiting further vision processing (e.g., estimation of normal and albedo maps). The presented evaluation provides the reader a broad view of the topic, which is of practical use for both researchers and practitioners. Our method can be easily and efficiently integrated into heterogeneous, existing pipelines, and even into web-services that take a raw MLIC and automatically return a particular defined outcome for visualization and relighting. Its integration in standard pipelines does not require substantial changes, and can be employed with a negligible training effort. The presented method has been applied to the classic MLIC

setup. Since the light direction and intensity calibrations are computed for each image independently, our method can be employed in other, different setup conditions, e.g., multi-view MLIC. In the future we will investigate how to adapt and exploit the proposed calibration solution to those more complex capture scenarios.

References

1. Ackermann, J., Fuhrmann, S., Goesele, M.: Geometric point light source calibration. In: Proc. VMV, pp. 161–168 (2013)
2. Alldrin, N.G., Mallick, S.P., Kriegman, D.J.: Resolving the generalized bas-relief ambiguity by entropy minimization. In: Proc. CVPR, pp. 1–7 (2007)
3. Angelopoulou, M.E., Petrou, M.: Uncalibrated flatfielding and illumination vector estimationfor photometric stereo face reconstruction. Machine vision and applications **25**(5), 1317–1332 (2014)
4. Barsky, S., Petrou, M.: The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. IEEE TPAMI **25**(10), 1239–1252 (2003)
5. Basri, R., Jacobs, D., Kemelmacher, I.: Photometric stereo with general, unknown lighting. International Journal of computer vision **72**(3), 239–257 (2007)
6. Chen, G., Han, K., Shi, B., Matsushita, Y., Wong, K.Y.K.: Self-calibrating deep photometric stereo networks. In: Proc. CVPR, pp. 8739–8747 (2019)
7. Chen, G., Han, K., Wong, K.Y.K.: PS-FCN: A flexible learning framework for photometric stereo. In: Proc. ECCV, pp. 3–18 (2018)
8. CHI: Cultural heritage imaging website (2020). URL http://culturalheritageimaging.org. [Online; accessed-March-2019]
9. Ciortan, I., Pintus, R., Marchioro, G., Daffara, C., Giachetti, A., Gobbetti, E.: A practical reflectance transformation imaging pipeline for surface characterization in cultural heritage. In: Proc. GCH, pp. 127–136 (2016)
10. Dong, Y., Chen, G., Peers, P., Zhang, J., Tong, X.: Appearance-from-motion: Recovering spatially varying surface reflectance under unknown lighting. ACM TOG **33**(6), 1–12 (2014)
11. Gardner, A., Tchou, C., Hawkins, T., Debevec, P.: Linear light source reflectometry. ACM TOG **22**(3), 749–758 (2003)
12. Georgiev, G.T., Butler, J.J.: BRDF study of gray-scale Spectralon. In: Earth Observing Systems XIII, vol. 7081, p. 708107. International Society for Optics and Photonics (2008)
13. Giachetti, A., Ciortan, I., Daffara, C., Marchioro, G., Pintus, R., Gobbetti, E.: A novel framework for highlight reflectance transformation imaging. Computer Vision and Image Understanding **168**, 118–131 (2018)
14. Habel, R., Wimmer, M.: Efficient irradiance normal mapping. In: Proc. I3D, pp. 189–195 (2010)
15. Huang, X., Walton, M., Bearman, G., Cossairt, O.: Near light correction for image relighting and 3D shape recovery. In: Proc. Digital Heritage, vol. 1, pp. 215–222 (2015)
16. Ikeuchi, K., Horn, B.K.: An application of the photometric stereo method. Tech. Rep. AI Memo 539, MIT AI Lab (1979)
17. Jung, J., Lee, J.Y., So Kweon, I.: One-day outdoor photometric stereo via skylight estimation. In: Proc. CVPR, pp. 4521–4529 (2015)
18. Koppal, S.J., Narasimhan, S.G.: Novel depth cues from uncalibrated near-field lighting. In: Proc. ICCV, pp. 1–8 (2007)
19. Li, B., Feng, J., Zhou, B.: A SVBRDF modeling pipeline using pixel clustering. arXiv preprint arXiv:1912.00321 (2019)
20. Lu, F., Matsushita, Y., Sato, I., Okabe, T., Sato, Y.: Uncalibrated photometric stereo for unknown isotropic reflectances. In: Proc. CVPR, pp. 1490–1497 (2013)
21. Ma, L., Liu, J., Pei, X., Hu, Y., Sun, F.: Calibration of position and orientation for point light source synchronously with single image in photometric stereo. Optics express **27**(4), 4024–4033 (2019)
22. 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)
23. Migita, T., Ogino, S., Shakunaga, T.: Direct bundle estimation for recovery of shape, reflectance property and light position. In: Proc. ECCV, pp. 412–425 (2008)
24. Nimier-David, M., Vicini, D., Zeltner, T., Jakob, W.: Mitsuba 2: A retargetable forward and inverse renderer. ACM TOG **38**(6), 1–17 (2019)
25. Papadhimitri, T., Favaro, P.: A closed-form, consistent and robust solution to uncalibrated photometric stereo via local diffuse reflectance maxima. International journal of computer vision **107**(2), 139–154 (2014)
26. Papadhimitri, T., Favaro, P.: Uncalibrated near-light photometric stereo. In: Proceedings of the British Machine Vision Conference, pp. 1–12. BMVA Press (2014)
27. Pintus, R., Ciortan, I., Giachetti, A., Gobbetti, E.: Practical free-form RTI acquisition with local spot lights. In: Proc. STAG, pp. 143–150 (2016)
28. Pintus, R., Dulache, T., Ciortan, I., Gobbetti, E., Giachetti, A.: State-of-the-art in multi-light image collections for surface visualization and analysis. Computer Graphics Forum **38**(3), 909–934 (2019)
29. 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**(3), 313–340 (2018)
30. Ramamoorthi, R., Hanrahan, P.: An efficient representation for irradiance environment maps. In: Proc. SIGGRAPH, pp. 497–500 (2001)
31. Sato, I., Okabe, T., Yu, Q., Sato, Y.: Shape reconstruction based on similarity in radiance changes under varying illumination. In: Proc. ICCV, pp. 1–8 (2007)
32. Sun, J., Smith, M., Smith, L., Farooq, A.: Sampling light field for photometric stereo. International Journal of Computer Theory and Engineering **5**(1), 14–18 (2013)
33. Wong, K.Y.K., Schnieders, D., Li, S.: Recovering light directions and camera poses from a single sphere. In: Proc. ECCV, pp. 631–642 (2008)
34. Woodham, R.J.: Photometric method for determining surface orientation from multiple images. Optical engineering **19**(1), 191,139 (1980)
35. 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)