

Practical Free-form RTI Acquisition with Local Spot Lights

R. Pintus¹, I. Ciortan², A. Giachetti², and E. Gobbetti¹

¹CRS4, Visual Computing Group, Italy

²University of Verona, Italy

Abstract

We present an automated light calibration pipeline for free-form acquisition of shape and reflectance of objects using common off-the-shelf illuminators, such as LED lights, that can be placed arbitrarily close to the objects. We acquire multiple digital photographs of the studied object shot from a stationary camera. In each photograph, a light is freely positioned around the object in order to cover a wide variety of illumination directions. While common free-form acquisition approaches are based on the simplifying assumptions that the light sources are either sufficiently far from the object that all incoming light can be modeled using parallel rays, or that lights are local points emitting uniformly in space, we use the more realistic model of a scene lit by a moving local spot light with exponential fall-off depending on the cosine of the angle between the spot light optical axis and the illumination direction, raised to the power of the spot exponent. We recover all spot light parameters using a multipass numerical method. First, light positions are determined using standard methods used in photometric stereo approaches. Then, we exploit measures taken on a Lambertian reference planar object to recover the spot light exponent and the per-image spot light optical axis; we minimize the difference between the observed reflectance and the reflectance synthesized by using the near-field Lambertian equation. The optimization is performed in two passes, first generating a starting solution and then refining it using a Levenberg-Marquardt iterative minimizer. We demonstrate the effectiveness of the method based on an error analysis performed on analytical datasets, as well as on real-world experiments.

Categories and Subject Descriptors (according to ACM CCS): I.4.1 [Image processing and Computer Vision]: Digitization and Image Capture—Imaging Geometry

1. Introduction

Multi-light reflectance processing techniques, such as Polynomial Texture Maps (PTM) [MGW01], more generic Reflectance Transformation Imaging (RTI) [MVSL05] and Photometric Stereo (PS) [Woo78, Woo80], aim to visually characterize objects by observing them from a fixed point of view under different lighting conditions. They are currently emerging as a de-facto standard in appearance and geometry acquisition due to their cost-effectiveness and flexibility. Their range of application goes from qualitative estimation of image formation models, for applications such as visual enhancement or relighting [MWGA06], to the quantitative recovery of shape and material properties [AG15].

While some techniques exist for recovering shape and material information under unknown environmental illumination conditions [BJK07, HWM*15], the most widespread approach used in many application fields, including Cultural Heritage (CH) investigation [MMSL06], medical interventions [DBO*15, DGL*14], and underwater data gathering [MCOC15], considers a single calibrated camera taking multiple images of a scene illuminated by a single moving light. A wide variety of physical realizations exist, ranging from purely free-form methods using a hand-held illuminator [CHI16], to different sizes of fixed light domes

[CHI16, SSWK13, Ham15]. While dome solutions allow for easier pre-calibration, and thus more reliable data, they are more costly and less flexible in terms of achievable object size and illumination directions. For this reason, much research is focusing on improving the quality of free-form solutions [GDR*15, CPM*16], in which a light is freely moved in front of a still camera.

Classic methods, however, assume for simplicity either a collimated and uniform light source (e.g., far point light), which means that the light direction and its intensity are the same across the entire image domain, or a local point light, which means that the only variation of illumination is due to the inverse-square distance fall-off. These assumptions do not hold for common off-the-shelf illuminators, such as LED lights, which, in addition to being placed arbitrarily close to the objects, present a variable angular radiation pattern [MS08]. Such simplifying assumptions induce considerable errors in the estimation of shape and material properties (see Sec. 2).

Our approach. In this paper, we present an automated light calibration pipeline for free-form acquisition of shape and reflectance of objects using common off-the-shelf illuminators, such as LED lights, that can be placed arbitrarily close to the objects. We overcome the limitations of common simplified lighting models, based

on parallel lighting or local point lights, by using a spot light model, which is general enough to model common illuminators such as LEDs [MS08]. The parameters that control lighting are thus the spot light position, the direction of the light optical axis, and a spot exponent that controls the exponential fall-off depending on the cosine of the angle between the spot light optical axis and the illumination direction. We recover all spot light parameters using a multipass numerical method. First, light positions are determined using standard methods used in photometric stereo approaches. Then, we exploit measures taken on a Lambertian reference planar object to recover the spot light exponent and the per-image spot light optical axis; we minimize the difference between the observed reflectance and the reflectance synthesized by using the near-field Lambertian equation. The optimization is performed in two passes, first generating a starting solution and then refining it using a Levenberg-Marquardt iterative minimizer.

Contribution. While not all the techniques presented in this work are novel in themselves, their elaboration and combination provide a significant step towards creating a practical and flexible RTI and PTM acquisition pipeline with local spot lights, such as the now ubiquitous LEDs. We introduce, in particular, a fully automated multipass numerical method for estimating the light parameters. The method is flexible enough to be applied to spot lights as well as to point light sources and Lambertian emitters. We demonstrate its effectiveness based on an error analysis performed on analytical datasets, as well as on real-world experiments.

Advantages and limitations. The proposed pipeline has several advantages. It is very easy to implement and to use and does not require any special object-dependent parameters to tune its behavior, so that it can be used by non-experts in application fields (e.g., CH, medical, etc). The proposed spot light model is general enough to model common illuminators such as LEDs, much better than common directional or point lighting models. Moreover, the light does not have to be placed far from the subject, but the light direction and intensity calibration can cope with close-range illumination. Our calibration and spotlight modeling techniques can be integrated to many existing pipelines, offering an improvement over those in common use, using simple directional lights or local point lights. In fact, it requires the same input data as the standard PS captures, and it is compliant to illuminators generally used. It is not only restricted to the case of free-form settings, but can be employed to calibrate fixed systems (e.g., dome-like light stages), or to test the quality and accuracy of robotic-based solutions. Of course, the main limitation is that the presence of some calibration target is required to compute the light position in space for each image. Although this excludes the application of our method to extreme wild setups, nowadays free-form PS acquisitions are employed daily by non-experts, which already feel comfortable to position calibration targets (reflective spheres and white reference targets) near the object under study before the capture session [CHI16].

2. Related work

The calibration of light direction and intensity for PS purposes is a wide and well known research subject. In the last decades a lot of methods have been published that deal with several types of illuminants, different from the classic far, point light condition, and with

their non-trivial calibration [PSG01, AFG13, TMNM09, WSL08]. The vast majority of them tries to calibrate near, point light emitters [ATS*12, WC01]. An exhaustive review of those methods is out of the scope of this paper, and we refer the reader to the survey of Ackermann and Goesele [AG15] both as a good, up-to-date review of the most important PS techniques, and as a useful source of information on error and calibration issues in this field. In the following, we provide only an analysis of the most recent and relevant works related to light calibration in PS, and we discuss the motivation of our contribution compared to them.

Some approaches try to calibrate light direction and intensity by interpolation means in the 2D image domain [GDR*15]. Ciortan et al. [CPM*16] estimate light directions in a small set of image locations by using reflective spheres (highlight pixels). They find a per-pixel light direction $l = \{l_x, l_y, l_z\}$ across the entire image by linear interpolating the sampled values of l_x and l_y components, and imposing $\|l\|^2 = 1$. They include a white planar target of known normal and albedo in the framed scene behind the captured object, and exploit the interpolated light direction information to first remove the Lambertian effect from measured radiance of white planar pixels, and then to compute light intensity for that region. Finally, they apply a quadratic interpolation to find light intensity values for the remaining part of the image domain. Their image-based computation does not take into account the effect due to a different depth of highlight points compared to the plane, and, although they do not impose a constraint on a light model, nonetheless they both force the light direction to follow a linear polynomial, and they make the assumption that the light intensity on the plane behaves as a quadratic function. Moreover, they do not explicitly take into account the fall-off due to the inverse of squared distance, by implicitly including it into the quadratic coefficients. Similarly, other methods [SSSF13, AP14] use a flat reference object with known albedo to calibrate an arbitrary lighting vector field. They don't use polynomial interpolation, but they exploit measured spatially-varying intensities to compensate the input images, and to convert the problem into a standard collimated case. Unfortunately, since these calibrations avoid to adopt a light model defined in the whole 3D space, they completely neglect direction and intensity variations in the z -axis. Hence, they are not generally applicable, being valid only in the vicinity of the calibration plane. Conversely, the proposed method relies on the camera internal calibration and the 3D calibration of target and light sources, i.e., white planar target equation in space and the 3D position of illuminators. This allows us to cast rays from/to the light sources and the camera, and to intersect them with the reference white plane. In this way, we are able to sample the light vector field (both direction and intensity) with a better accuracy along the planar reference, by taking also into account the fall-off due to the distance. Moreover, by adopting a defined illumination model, we know the light behavior in the 3D space, so that we can compute light directions and intensities for points outside the calibration plane.

Considering the light form factor and a model of its behavior in the 3D space for PS analysis is a very old and well known topic; for instance, the seminal approach by Ikeuchi and Horn [IH79] models the image formation process given a linear light source and a specular object. More recently, Mecca et al. [MWBK14] proposed a mathematical formulation based on quasi-linear PDEs, which

solves a perspective, near field photometric stereo given point light sources. Beside directional lighting, Quéau and Durou [QD15] show how to derive lighting models for several real-world scenarios, such as isotropic, near punctual model, extended light sources, and LCD-screen based illuminants. Although these works pose the mathematical bases for dealing with more general lighting conditions, however, the inverse calibration problem remains challenging, i.e., computing lighting model parameters from a given set of intensity measurements.

Two recent approaches are very similar to ours, and worth mentioning and comparing. Huang et al. [HWBC15] propose an alternating minimization formulation for computing near-light positions and surface normals. Their model takes into account several non-idealities, such as fall-off due to the squared distance and camera vignetting. Although they are capable to calibrate the system without knowing the light positions in advance, however, they impose a point light model, which limits the method applicability. Although we require that some (at least three) reflective spheres were included within the framed scene to initially estimate light positions, we assume that the form factor of the light source can be generally modeled by the product of an intensity and a cosine-power term. This is similar to the model presented by Xie et al [XSJ*15], which propose a LED-based photometric stereo system. They calibrate a seven LED lamp setup through the estimation of lighting positions and principal axis. They use a LED lighting model and a diffuse sphere of known shape as a calibration target; the sphere exhibits also a small specular signal. Unfortunately, they rely on a fixed light rig (not free-form), and on the fact that the optical axis of the LED is collinear to the incident light direction at a specular point on the sphere, which is not always the case, and, in addition, very rarely met in a free-form acquisition. Moreover, their approach depends on a fixed LED light, whose parameters are known a-priori. Instead, we consider the light intensity, its optical axis and the exponential as unknown variables. Compared to these similar techniques, in both cases our model allows us to deal with a free-form PS acquisition and a wide range of light sources, including ideal point lights, Lambertian emitters and the more general spot light illuminators.

3. Method

The main input of the proposed calibration pipeline is a set of images taken from the same view point but with different illumination conditions, together with the intrinsic parameters of the camera used for capture. The images are considered already undistorted, so we discard distortion coefficients, and we use only the camera matrix for our computation (i.e., focal lengths and principal point). We also require that a white Lambertian planar target were included within the framed scene; we need as input the corresponding plane equation in the camera reference frame, and a binary mask image that indicates the pixels that belong to that reference flat object. Further, the per-image light positions in the camera reference frame must be estimated prior to the application of our method. Such information can be obtained by a variety of standard means [AG15]. In this work, we derive it by using a calibrated camera combined with reflective spheres of known size. Similarly to Powell et al. [PSG01], sphere positions are detected in the im-

ages. Camera rays are then shot towards the observed highlight pixels, and the reflected rays are then intersected in 3D to find the light position for a given image. The rest of the method is, however, independent from the technique used to find the light position.

3.1. Lighting model and unknowns

Consider the region W of pixels $w = (u, v)$ belonging to the planar target. This region is the same for all images, since the PS/RTI setup requires a fixed view point. We model the formation process of image i by using the near-field Lambertian equation [HWBC15]:

$$I(i, w) = \rho(w) \frac{L(i, w)}{d(i, w)^2} (\hat{l}(i, w) \cdot \hat{n}(w)) \quad (1)$$

Each pixel with brightness I corresponds to a 3D point P in the target with albedo ρ and unit normal n ; d is the distance between the point and the light source position $P_L(i)$, while $\hat{l}(i, w) = (P_L(i) - P(w)) / \|P_L(i) - P(w)\|$ is the light direction at point $P(w)$. L is the emitted light intensity along the direction \hat{l} . We make the assumption that L is a function of the light optical axis \hat{a} and the specific light direction \hat{l} , by using the cosine-power term [XSJ*15]:

$$L(i, w) = L_0 (\hat{l}(i, w) \cdot \hat{a}(i))^m \quad (2)$$

where L_0 is the maximum emitted intensity (i.e., the emitted light intensity along the optical axis \hat{a}) and m is the exponential that regulates the intensity fall-off due to the angle between the light direction \hat{l} and its optical axis \hat{a} .

The provided mask defines the set of pixels W for all images corresponding to the white planar target, and the input set of images in that region gives us the per-pixel measured values of I . We consider $\rho = 1$ for the white target, and we know from its equation the value of \hat{n} ; this quantities are independent from w . By exploiting camera intrinsic parameters we can cast rays from the camera to the plane, and we can compute for each pixel the corresponding 3D point $P(w)$. Given the known light position $P_L(i)$ for each image i , it is straightforward to compute both $d(i, w)$ and $\hat{l}(i, w)$. Given this a-priori knowledge we compute from equation 1 the measured value of $\tilde{L}(i, w)$ as:

$$\tilde{L}(i, w) = \frac{I(i, w) d(i, w)^2}{\hat{l}(i, w) \cdot \hat{n}} \quad (3)$$

For an input image set of cardinality N , we have to find $N + 2$ unknown variables, i.e., L_0 , m and $\hat{a}(i)$ for $i = 1 \dots N$. Here we propose to solve this problem through a non-linear minimization strategy:

$$\underset{L_0, m, \hat{a}(i)}{\operatorname{argmin}} \sum_{i=1}^N \sum_w \|\tilde{L}(i, w) - L(i, w)\|^2 \quad (4)$$

3.2. Spot light calibration

In order to estimate the spot light parameters, we solve the global non-linear optimization problem using a two-step approach, in which a coarse global solution is obtained with a simplified model and then refined by a local minimization of a more complex objective function.

The first step is thus to find a reliable starting point for the search

of the local minimum of the sum of squares of nonlinear functions defined in equation 4. For each image we need to estimate the initial guess for $\hat{a}(i)$, and then we need to find a global initial value both for L_0 and m .

In our approach, in order to find the coarse initial solution, we first separately solve the problem in parallel for each single image, then merge the results. For each single image i , we follow a brute force approach; for simplicity, since we are dealing here with only one image, we omit the i in the description. For a fixed subset S of pixels in the image (see section 3.3 for details) we cast rays from the camera center of projection to the scene, and we compute the intersections of those rays and the plane defined by the planar target. For each of those 3D point $P(s)$ we define a candidate axis $\hat{a}(s) = (P_L - P(s)) / \|P_L - P(s)\|$, where P_L is the position of the light for this specific image. Note that S is a subset of the entire image, and it is not correlated to the set W ; this is because, for the sake of generality, the optical axis of the light might fall outside the region W . Conversely, at least for the initial guess, we assume that it falls within the framed scene. To find a pair $\{L_0(\hat{a}(s)), m(\hat{a}(s))\}$ for each candidate axis we use the logarithmic least-squares fitting

$$y = A + B \ln(x) \quad (5)$$

applied to equation 2, where $y = y(w) = \ln(L(w))$, $A = \ln(L_0(\hat{a}(s)))$, $B = m(\hat{a}(s))$, and $x = x(\hat{a}(s), w) = \hat{l}(w) \cdot \hat{a}(s)$. Considering all the N_W image pixels in W , it leads to:

$$\begin{cases} B = m = \frac{N_W \sum_w (y_w \ln x_w) - \sum_w y_w \sum_w \ln x_w}{N_W \sum_w (\ln x_w)^2 - (\sum_w \ln x_w)^2} \\ A = \ln(L_0) = \frac{\sum_w y_w - B \sum_w (\ln x_w)}{N_W} \end{cases} \quad (6)$$

For each candidate axis we compute the residual of the logarithmic fitting:

$$R = \sum_w \|A + B \ln(x_w) - y_w\|^2 \quad (7)$$

We initialize the light optical axis for the selected image with the value $\hat{a}(s)$ corresponding to the minimum residual, and we keep both the residual and the associated pair $\{L_0(\hat{a}(s)), m(\hat{a}(s))\}$ for further processing.

After we loop for all the images we obtain the initial assignment of all $\hat{a}(i)$ for $i = 1 \dots N$. Now we need to find the initial values of L_0 and m , which are common to all the input images. From the previous computation, for each image i , we keep the minimum residual (we call it now $R(i)$) and the corresponding intensity and exponential estimations (we call them now $\{L_0(i), m(i)\}$). We first order these pairs by increasing residual and, to discard possible outliers, we keep only pairs with residual values less than the median residual. For the remaining pairs, we choose the one that minimizes the sum of squared residuals in equation 7 across the entire input image set.

Finally, starting from the coarse solution just computed, we refine calibration by a local minimization of the logarithmic version of equation 4, globally computed within the set W across all input images, i.e.:

$$\underset{L_0, m, \hat{a}(i)}{\operatorname{argmin}} \sum_{i=1}^N \sum_w \| \ln \tilde{L}(i, w) - \ln L(i, w) \|^2 \quad (8)$$

This local nonlinear minimization problem is efficiently solved with the Levenberg-Marquardt optimization algorithm [LA04].

3.3. Implementation notes

As we mentioned above, the calibration pipeline is completely automatic. It does not need any user-defined parameter tuning that depends on the input data nature, and, in particular, it is independent from pixel resolution and number of light positions. However, some hidden parameters must be taken under control in order to make the pipeline scalable. More specifically, the cardinality of the sets W and S might strongly affect the computational time, so we need a strategy to keep those under a certain threshold. Thus, the algorithm will not consider the entire pixels in the mask (W), nor the entire image when computing candidate axis (S). In both cases, it automatically performs an uniform sub-sampling over these domains, by automatically setting two hidden parameters. The number of white planar pixels (cardinality of W) influences the speed of the Levenberg-Marquardt optimization algorithm; we drive the sampling over the masked region across all images so that the number of measurements in the optimization will not exceed 10^5 . Similarly, for each image we will extract no more than 10^4 candidate axes (cardinality of S). These two parameters are fixed once and never changed whatsoever. Further, to avoid effects due to a grid-like sub-sampling of pixels in those regions, in both cases we use a two-dimensional low-discrepancy picking sequence (Halton sequence [KN12]).

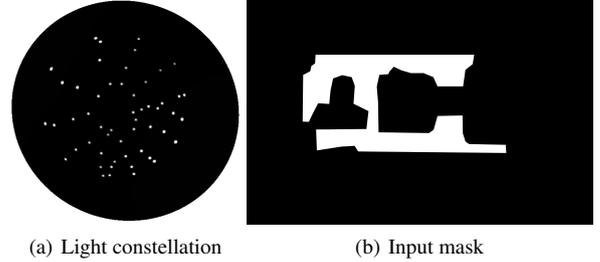


Figure 1: Synthetic setup. We use a controlled synthetic setup to validate light calibration accuracy. The light positions are taken from one real-world acquisition, and are shown here as a constellation of highlights on one of the sphere (a). The mask used to select white planar target pixels (b) has been created with a non-regular pattern (e.g., a simple rectangle) to simulate a real scenario with real objects occluding the planar target.

4. Results

The proposed light calibration algorithm has been tested both on synthetic and real datasets. The pipeline has been implemented on Linux using C++ and the OpenCV library [Ope13]. We employ as optimization code the Lourakis's open source C++ library for solving non-linear problem by Levenberg-Marquardt algorithm [Lou04]. Our benchmarks were executed on a PC with 4 Intel Core i7-4510U CPU @ 2.00GHz processors, and 8GB RAM.

In order to test and validate the quality of the proposed light

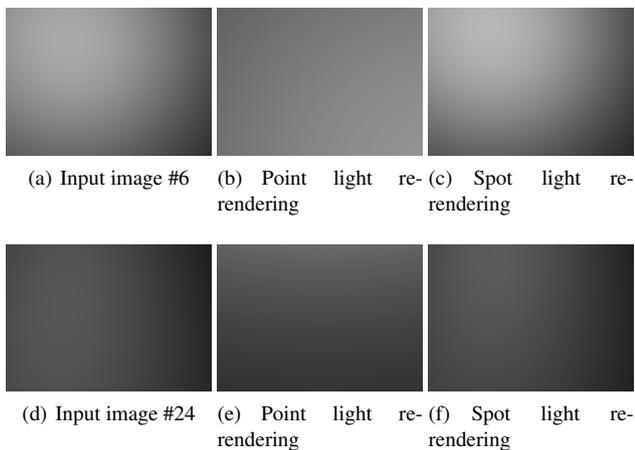


Figure 2: Re-rendered synthetic images. We take two input original images, #6 (a) and #23 (d), and we compare them with the re-rendering of the diffuse plane done with the light parameters computed by two different calibration models, i.e., point light (b) (e) and spot light (c) (f). Spot light model behaves better in fitting typical common LED-based emitters.

Method	Min	Max	Average	Median	Std
Point light	0.0	0.24	0.07	0.07	0.05
Spot light	0.0	0.06	0.02	0.01	0.01

Table 1: Average re-rendering error statistics.

calibration pipeline, we present here results obtained on synthetic data as well as a real-world acquisition. For the generation of all synthetic examples, we provide as input the same camera intrinsic parameters, and 53 light positions arranged in the hemisphere above the scene; figure 1(a) shows the light constellation as seen from one of the glossy spheres. We didn't compute the light constellation analytically, but instead we took it from one free-form real acquisition, in order to consider a reasonable real-world distribution. The size of the acquired field and the distance between the scene and the light sources are the same as common RTI capture setup. Accordingly, we set illumination parameters similar to the most common LED light emitters. As in the paper by Xie et al. [XSJ*15], we chose a LED with an emitting angle of $\theta = 15^\circ$; this means that the light exponential is $m = -\ln 2 / \ln(\cos \theta) \approx 20$. Since the equation 1 includes the inverse of the squared distance, light intensity value is set to $L_0 = 522000$ in order to produce a meaningful average exposure value across the rendered images under the chosen light constellation. The light axes are randomly set within those that fall on the visual portion of the plane; in other words, no light axis intersects the plane in a point not visible from the camera. Finally the input mask is created with a non-regular pattern in order to simulate objects positioned within the scene that occlude the planar white target (Figure 1(b)). In all the experiments images the signal is a grayscale brightness, and its range is normalized to $[0, 1]$.

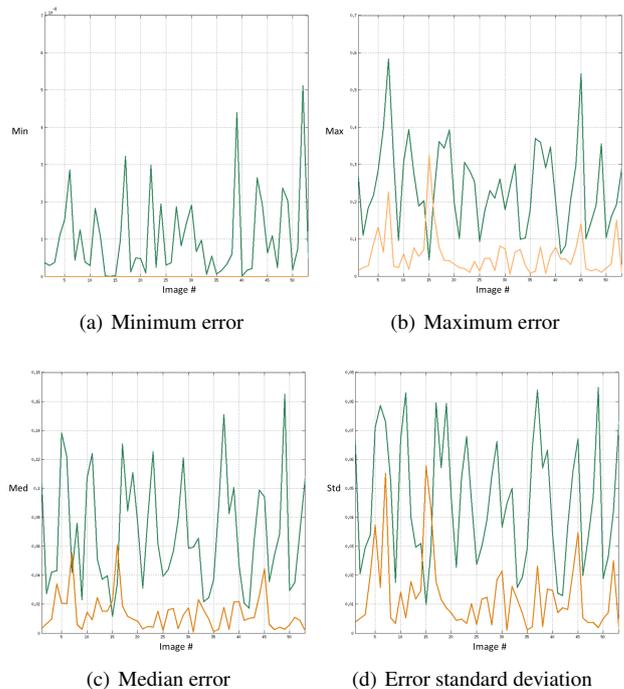


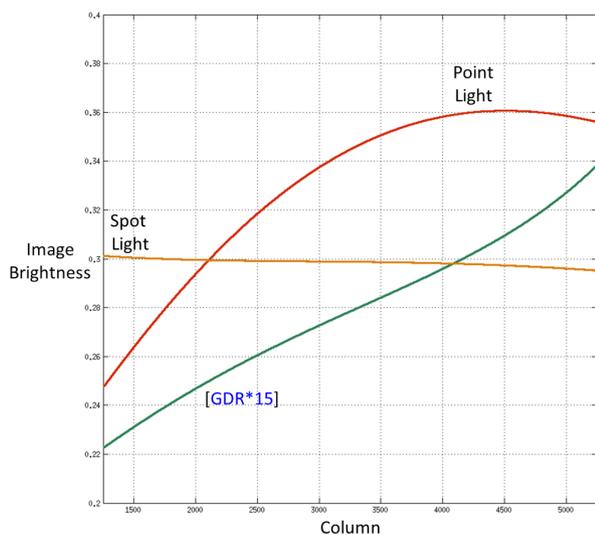
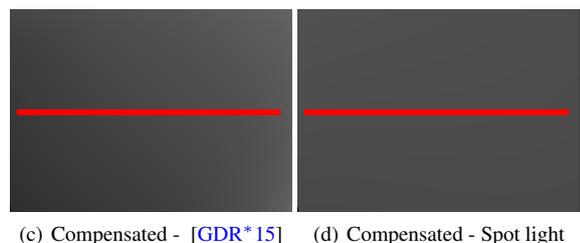
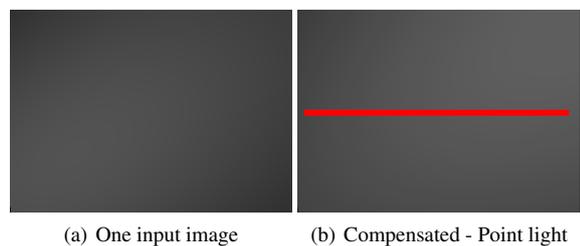
Figure 3: Re-rendering error plots. We consider the pixels belonging to the diffuse plane and we plot the per-image minimum (a), maximum (b), and median (c) re-rendering error, and its corresponding standard deviation (d). The spot light re-rendering (orange lines) exhibits a much lower error than the point light one (green lines).

Method	Min	Max	Average	Median	Std
Point light	0.03	5.3	2.6	2.7	1.3
[GDR*15]	0.04	13.1	6.3	6.1	2.9
Spot light	0.03	3.1	1.6	1.6	0.6

Table 2: Normal angle error statistics (degrees).

4.1. Synthetic plane

We first test our approach on a controlled, simple scenario, i.e., a synthetic diffuse plane. The plane is rendered by using the near-field Lambertian model in equation 1. We first test the difference between the ideal point light assumption and the calibration of a spot light. In the first case, only light positions are required, and we need to find only an estimation of the light intensity, in order to take into account fall-off due to the inverse of squared distance. In the case of spot light, light intensity and positions are not enough, and we require also light optical axis and decay factor. To measure the quality of the two assumptions, we re-render the images of the diffuse plane by using the parameters computed in the two calibration approaches (point light vs spot light) and the equation 1. In Figure 2 we show two examples. For each of them we present the original image, and the two diffuse planes re-rendered by using point light and spot light model. In the first row the simulated light is directly above the scene (see highlights in the spheres) while, in



(e) Single row plot of compensated signals

Figure 4: Compensated signal. We give the same synthetic planar dataset to the light calibration pipeline based on point light source, 2D interpolation [GDR*15], and to our algorithm. We show here one of the input image (a), and the three respective compensated images after the light calibration (b) (c) (d). Since the framed object is a plane, the compensated image should have a flat, uniform appearance. We compare the signal of the three methods by plotting a selected part of a single row 4(e). This result measures how our method exhibits a higher capacity in compensating non-uniform lighting.

the bottom row, we have a more raking light. In both cases the light optical axis points toward the central part of the image. It is clear how the spot light model fits better the simulated LED illuminant. To compare them numerically, we compute some error statistics in the two cases of point light or spot light re-rendering. In Figure 3

we plot the per-image minimum 3(a), maximum 3(b), and median error 3(c), and the corresponding standard deviation 3(d). The spot light calibration (orange lines) clearly outperforms the point light one (green lines). For the sake of readability the y-axes ranges of the plots are all different. In table 1 we present the average values of the same statistics computed across all images. Here we also show how the average and the median are almost equal; the same holds for the per-image values, so that we omit the corresponding plot in Figure 3.

In Figure 4 we show one of the original input images from the synthetic planar dataset. After the light calibration, the knowledge of light direction and intensity allows for the computation of a corrected compensated image, as if the light was a perfect uniform point light emitter. We compare the quality of calibration by computing this image with the three methods based on point light model, quadratic light intensity interpolation [GDR*15] and spot light model. What we expect is that, since the framed object is a plane, the higher is the accuracy of light calibration, the flatter will be the signal in the compensated image. Figure 4(b), 4(c) and 4(d) respectively show these compensated images. In order to make the result more readable, we selected a part of a row in these three images (red lines), and we plot only that signal (Figure 4(e)). Our pixel values are flatter than those found with the other light calibration pipelines, so spot light model allows for a better counterbalance of the non-homogeneous light intensity across the image.

Finally we compared the plane normal computation by applying photometric stereo to the three calibrated image sets. Since we already know the normal of the synthetic plane, we computed the angle error statistics produced by point light model, image based interpolation and spot light assumption. The typical quality of a normal estimation in PS algorithm is in the range [2,5] degree error [AG15]. We can see how our approach is in line with state-of-the-art baseline accuracy, and performs better than the point light based calibration and the image-based approaches [GDR*15].

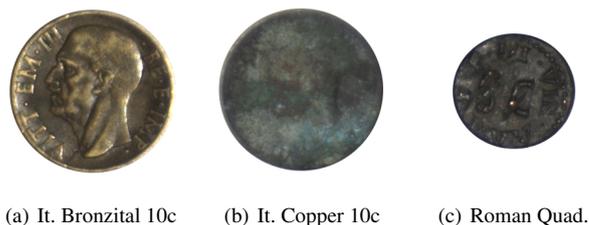


Figure 5: Samples. Photos of the real-world coins used to test the proposed light calibration approach: (a) an Italian Bronzital 10c coin; (b) a severely degraded Italian copper 10c coin; (c) a bronze roman coin.

4.2. Physical acquisition results

In order to have an early validation of our method in a practical setting, we present here results obtained for the acquisition of a

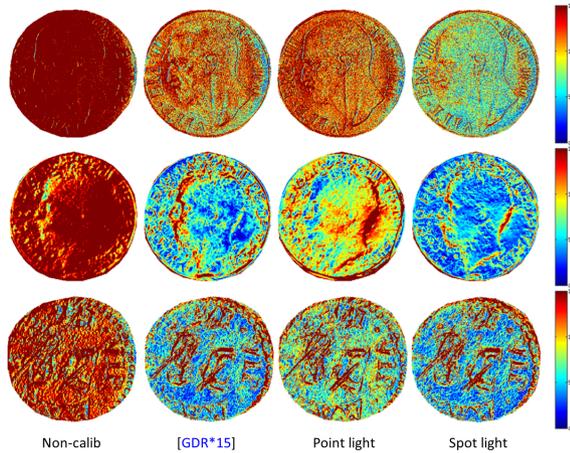


Figure 6: Normal map errors. Difference between PS estimated normals from the reference microprofilometer normals.

Table 3: Median angular distances of the PS estimated normals from the reference microprofilometer normals.

	Non-calib.	[GDR*15]	Point	Spot
Bronzital 10c	15.08	5.95	7.39	4.93
Copper 10c	20.62	11.41	12.29	8.06
Quadrans	13.84	6.82	8.38	6.68

small bronze Roman coin (*quadrans*) and two 10 cent antique Italian coins (see Figure 5). The Roman coin, dated 9 BC, is made of bronze and damaged by scratches. One exemplar of the 10c Italian coins, dated 1931 is made of copper and is severely degraded, while the second one, dated 1939, is made of a special alloy with nickel called Bronzital, which has been used to improve corrosion resistance.

The digital camera used for the experiment is a DSLR Nikon D810 with CMOS sensor (36x24mm, spatial resolution of 36MP for the full format image area), with a AF-S FX Nikkor 50mm f/1.8G lens and was remote-controlled to avoid vibrations. The light source used for the experiment was a white LED (color temperature 6500K). All images were acquired at ISO 32, aperture f8, shutter speed 0.8s, with custom white balance and manual focus. The distance from the camera to the acquisition plan was approximately 95cm, while the light was positioned at a distance of about 1m. from the target object. 49 light directions were acquired around the object. The position of the lights and of the calibration plane was estimated from highlights in four reference spheres of known size placed at the corners of the acquisition area.

In order to test the accuracy of shape capturing, we compared a PS reconstruction of the normal field with measures of the same coins taken with an optical microprofilometer with a transversal resolution (XY grid) of 50 microns with a repeatability (Z heights) of 0.1 microns. Normals were reconstructed without any correction, as done in many RTI approaches for relighting or shape enhancement applications, as well the in-plane calibration using low-degree polynomials [GDR*15], a point light model, and our spot light

model. Normal maps of the photographic and microprofilometric acquisitions were registered with a mutual information maximization approach after a manual pre-alignment.

As shown in Table 3 and Figure 6, light calibration sensibly improves normal quality, and the spot light model consistently provides measurable improvements with respect to the other approaches.

5. Conclusion and future work

We have presented an automated light calibration pipeline for achieving accurate free-form RTI acquisition using common off-the-shelf illuminators, such as LED lights, that can be placed arbitrarily close to the objects and can loosely point at them. Reconstruction is achieved by modeling the moving illuminator as a spot light, and recovering the parameters through a multipass numerical method that optimizes the difference between the observed reflectance on a reference white planar target and the reflectance synthesized by using the near-field Lambertian equation with a spot light illuminator.

Our results on synthetic and real data demonstrate that this model produces measurably more accurate results in normal reconstruction than competing methods based on local point light modeling or in-plane low-frequency lighting correction.

In order to further improve its accuracy and light modeling capability, we plan to further extend our method to estimate a multiple-lobe radiation pattern for the moving light [MS08]. We also plan to use this technique in material characterization applications.

Acknowledgments. This work was partially supported by the Scan4Reco project funded by European Union's Horizon 2020 Framework Programme for Research and Innovation under grant agreement no 665091. We also acknowledge the contribution of Sardinian Regional Authorities under projects VIGEC and Vis&VideoLab.

References

- [AFG13] ACKERMANN J., FUHRMANN S., GOESELE M.: Geometric point light source calibration. In *VMV* (2013), pp. 161–168. 2
- [AG15] ACKERMANN J., GOESELE M.: A survey of photometric stereo techniques. *Foundations and Trends® in Computer Graphics and Vision* 9, 3-4 (2015), 149–254. 1, 2, 3, 6
- [AP14] ANGELOPOULOU M. E., PETROU M.: Uncalibrated flatfielding and illumination vector estimation for photometric stereo face reconstruction. *Machine vision and applications* 25, 5 (2014), 1317–1332. 2
- [ATS*12] AOTO T., TAKETOMI T., SATO T., MUKAIGAWA Y., YOKOYA N.: Position estimation of near point light sources using a clear hollow sphere. In *ICPR* (2012), pp. 3721–3724. 2
- [BJK07] BASRI R., JACOBS D., KEMELMACHER I.: Photometric stereo with general, unknown lighting. *International Journal of Computer Vision* 72, 3 (2007), 239–257. 1
- [CHI16] CHI: Cultural heritage imaging, 2016. URL: <http://culturalheritageimaging.org>. 1, 2
- [CPM*16] 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 *The 14th Eurographics Workshop on Graphics and Cultural Heritage* (October 2016). To appear. 1, 2

- [DBO*15] DULIU A., BROSIG R., OGNAWALA S., LASSER T., ZIAI M., NAVAB N.: Illumination compensation and normalization using low-rank decomposition of multispectral images in dermatology. In *International Conference on Information Processing in Medical Imaging* (2015), Springer, pp. 613–625. 1
- [DGL*14] DURR N. J., GONZÁLEZ G., LIM D., TRAVERSO G., NISHIOKA N. S., VAKOC B. J., PAROT V.: System for clinical photometric stereo endoscopy. In *SPIE BiOS* (2014), International Society for Optics and Photonics, pp. 89351F–89351F. 1
- [GDR*15] GIACHETTI A., DAFFARA C., REGHELIN C., GOBBETTI E., PINTUS R.: Light calibration and quality assessment methods for reflectance transformation imaging applied to artworks' analysis. In *SPIE Optical Metrology* (2015), International Society for Optics and Photonics, pp. 95270B–95270B. 1, 2, 5, 6, 7
- [Ham15] HAMEEUW H.: Mesopotamian clay cones in the ancient near east collections of the royal museums of art and history. *Bulletin van de Koninklijke Musea voor Kunst en Geschiedenis* 84 (2015), 5–48. 1
- [HWBC15] HUANG X., WALTON M., BEARMAN G., COSSAIRT O.: Near light correction for image relighting and 3d shape recovery. In *2015 Digital Heritage* (2015), vol. 1, IEEE, pp. 215–222. 3
- [HWM*15] HUNG C.-H., WU T.-P., MATSUSHITA Y., XU L., JIA J., TANG C.-K.: Photometric stereo in the wild. In *2015 IEEE Winter Conference on Applications of Computer Vision* (2015), IEEE, pp. 302–309. 1
- [IH79] IKEUCHI K., HORN B. K.: *An Application of the Photometric Stereo Method*. Tech. rep., DTIC Document, 1979. 2
- [KN12] KUIPERS L., NIEDERREITER H.: *Uniform distribution of sequences*. Courier Corporation, 2012. 4
- [LA04] LOURAKIS M., ARGYROS A.: *The design and implementation of a generic sparse bundle adjustment software package based on the levenberg-marquardt algorithm*. Tech. rep., Institute of Computer Science-FORTH, Heraklion, Crete, Greece, 2004. 4
- [Lou04] LOURAKIS M.: levmar: Levenberg-marquardt nonlinear least squares algorithms in C/C++. <http://www.ics.forth.gr/lourakis/levmar/>, Jul. 2004. [Accessed on 31 Jan. 2005.]. 4
- [MCO15] MASSOT-CAMPOS M., OLIVER-CODINA G.: Optical sensors and methods for underwater 3d reconstruction. *Sensors* 15, 12 (2015), 31525–31557. 1
- [MGW01] MALZBENDER T., GELB D., WOLTERS H.: Polynomial texture maps. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques* (2001), ACM, pp. 519–528. 1
- [MMSL06] MUDGE M., MALZBENDER T., SCHROER C., LUM M.: New reflection transformation imaging methods for rock art and multiple-viewpoint display. In *VAST* (2006), vol. 6, Citeseer, pp. 195–202. 1
- [MS08] MORENO I., SUN C.-C.: Modeling the radiation pattern of leds. *Opt. Express* 16, 3 (Feb 2008), 1808–1819. 1, 2, 7
- [MVSL05] MUDGE M., VOUTAZ J.-P., SCHROER C., LUM M.: Reflection transformation imaging and virtual representations of coins from the hospice of the grand st. bernard. In *VAST* (2005), vol. 2005, p. 6th. 1
- [MWBK14] MECCA R., WETZLER A., BRUCKSTEIN A. M., KIMMEL R.: Near field photometric stereo with point light sources. *SIAM Journal on Imaging Sciences* 7, 4 (2014), 2732–2770. 2
- [MWGA06] MALZBENDER T., WILBURN B., GELB D., AMBRISCO B.: Surface enhancement using real-time photometric stereo and reflectance transformation. *Rendering techniques 2006* (2006), 17th. 1
- [Ope13] OPENCV: Opencv - open source computer vision library, 2013. URL: <http://opencv.org/>. 4
- [PSG01] POWELL M. W., SARKAR S., GOLDFOF D.: A simple strategy for calibrating the geometry of light sources. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 9 (2001), 1022–1027. 2, 3
- [QD15] QUÉAU Y., DUROU J.-D.: Some illumination models for industrial applications of photometric stereo. In *The International Conference on Quality Control by Artificial Vision 2015* (2015), International Society for Optics and Photonics, pp. 95341C–95341C. 3
- [SSSF13] SUN J., SMITH M., SMITH L., FAROOQ A.: Sampling light field for photometric stereo. *International Journal of Computer Theory and Engineering* 5, 1 (2013), 14. 2
- [SSWK13] SCHWARTZ C., SARLETTE R., WEINMANN M., KLEIN R.: Dome ii: A parallelized btf acquisition system. In *Proceedings of the Eurographics 2013 Workshop on Material Appearance Modeling: Issues and Acquisition* (2013), Eurographics Association, pp. 25–31. 1
- [TMNM09] TAKAI T., MAKI A., NIINUMA K., MATSUYAMA T.: Difference sphere: an approach to near light source estimation. *Computer Vision and Image Understanding* 113, 9 (2009), 966–978. 2
- [WC01] WEBER M., CIPOLLA R.: A practical method for estimation of point light-sources. In *BMVC* (2001), pp. 1–10. 2
- [Woo78] WOODHAM R. J.: *Photometric Stereo*. Tech. rep., DTIC Document, 1978. 1
- [Woo80] WOODHAM R. J.: Photometric method for determining surface orientation from multiple images. *Optical engineering* 19, 1 (1980), 191139–191139. 1
- [WSL08] WONG K.-Y. K., SCHNIEDERS D., LI S.: Recovering light directions and camera poses from a single sphere. In *European conference on computer vision* (2008), Springer, pp. 631–642. 2
- [XSJ*15] 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 (2015), 42–50. 3, 5