EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2016)
C. E. Catalano and L. De Luca (Editors)

A Practical Reflectance Transformation Imaging Pipeline for Surface Characterization in Cultural Heritage

I. Ciortan[1], R. Pintus[2], G. Marchioro[1], C. Daffara[1], A. Giachetti[1], and E. Gobbetti[2],

[1]University of Verona, Italy
[2]CRS4, Visual Computing Group, Italy

Abstract

We present a practical acquisition and processing pipeline to characterize the surface structure of cultural heritage objects. Using a free-form Reflectance Transformation Imaging (RTI) approach, 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. Multiple reflective spheres and white Lambertian surfaces are added to the scene to automatically recover light positions and to compensate for non-uniform illumination. An estimation of geometry and reflectance parameters (e.g., albedo, normals, polynomial texture maps coefficients) is then performed to locally characterize surface properties. The resulting object description is stable and representative enough of surface features to reliably provide a characterization of measured surfaces. We validate our approach by comparing RTI-acquired data with data acquired with a high-resolution microprofilometer.

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

1. Introduction

The study of materials used in cultural objects is of extreme importance for the understanding and preservation of Cultural Heritage (CH). Curators, restorers, conservators and conservation scientists routinely exploit information on materials to have the proper insights about the time and way an artwork has been made, the techniques used, the environmental conditions of preservation, the previous conservation interventions, as well as to gather indications for planning future interventions. Materials in CH are studied under many different angles, ranging from the acquisition of basic information, such as surface average color and roughness, to the analysis of their molecular and elemental components. State-of-the-art study and conservation practices thus combine a wide variety of measurement probes and analytical techniques. Moreover, in order to study dynamic processes, such as the effects of aging, weathering, and restoration treatments, laboratory studies are often performed on appropriately prepared samples and mock-ups.

Among the different material characterizations, the study of surface appearance, in terms of reflectance and geometric meso- and micro-structure is of particular importance, since the vast majority of the cultural information is conveyed through optical signals from the viewed artwork to the human vision system. Characterizing surface structure and appearance is thus paramount for a variety of CH applications, from the assessment of the visual effects of restoration treatments, to the high-fidelity virtual and physical replication of cultural objects trough graphics and fabrication means.

Reflectance Transformation Imaging (RTI) [Woo78, MGW01, MVSL05], which derives material and geometry information from photographs of a fixed object under fixed camera position with varying lighting conditions, are nowadays a de-facto standard in appearance and geometry acquisition, due to their cost-effectiveness and flexibility. Light lies at the core of the artistic process, and illumination under multiple light conditions and directions is known to have been used by artists throughout time in order to obtain a more holistic representation when depicting a composition [PCC*10]. Correspondingly, in order to decipher an artistic technique or to understand the underlying information about an artifact, multi-light imaging at raking angles [JMS*14] provides a fair solution for the documentation, recording and decoding of CH objects [MMSL06].

Using RTI for surface characterization requires, however, to go beyond purely qualitative measurements. While completely calibrated setups, based on fixed domes, do exist, they have considerable costs and limitations in terms of scalability to multiple object sizes and achievable illumination configurations. The most practical and flexible techniques are based, instead, on the combination of freely moving lights with calibration objects to recover illumination directions (see Sec. 2). However, due to the lack of uniformity in illumination intensity and direction, the results obtained with this simple setup may vary widely between acquisitions, and may be not

suitable for quantitative analyses, which include normal estimation, roughness or material segmentation/classification, and monitoring over time. Exploitation of RTI data is thus limited.

Our approach. We introduce a simple and effective pipeline for the acquisition of geometry or visual appearance of cultural heritage objects in the case of near-light condition and spatially varying illumination. Our setup consists of a DSLR camera, a hand held light, and several calibration target, i.e., a set of glossy dark spheres and a white background (see Sec. 3). The algorithm takes an image stack containing multiple photos taken while light is projected from different directions. In the calibration step the images are undistorted. Then, the pipeline exploits sphere targets to cope with a near-light condition, where the parallel ray assumption is violated and a per-pixel light direction is estimated. The white target helps to deal with a non-point and non-uniform light, by balancing the scene irradiance without explicitly requiring the light source position in space to correct the light intensity as a function of the inverse squared distance, not to mention its more complex form factor (see Sec. 3). After that a corrected scene radiance is taken per-pixel in the form of an appearance profile [KN06]. An estimation of geometry and eventually reflectance parameters (e.g., albedo, normals, polynomial texture maps coefficients) is then performed to locally characterize surface and appearance properties. The setup and approach have been validated by comparing our results with those obtained with a high-resolution laser microprofilometer.

Contribution. While not all the techniques presented in this work are novel in themselves, their elaboration and combination provide a significant step towards a practical quantitative analysis using simple and low-cost uncalibrated acquisition strategies. We introduce, in particular, a setup and processing strategy that can cope with the effects of non-point lights, spatially varying illumination, and variable light distance, and we demonstrate the effectiveness of the approach even in challenging situations, such as damaged coins and silver samples used for restoration studies.

Advantages and limitations. Our approach is designed with the idea of being cost-effective, flexible, and simple to use. It offers a better reliability than current free-form approaches, while not requiring a fixed geometry, such as domes or robotic acquisitions, and/or complex calibration procedures. Since a single movable lamp is used, it is also easier and less costly than with multi-light setups to extend it to a multi- or hyper-spectral capture, since this could be achieved by just using narrow band illuminations or Liquid Crystal Tunable Filters (LCTF) in front of a monochromatic camera. The results presented here are currently limited to the capture of geometry and appearance by measuring low-frequency components of BRDFs, and must me extended to fully recover appearances of surfaces with complex reflectivity. This is, however, orthogonal to the approach presented in this paper.

2. Related work

Measurement of surface properties from multi-light imaging is a wide research subject. In the following, we provide an analysis of the most closely related approaches. For more details, we refer the reader to established surveys in digital material modeling [DRS10] and geometry and reflectance acquisition [AG15, WK15].

RTI setups. RTI transforms a high amount of visual information (2D image sets) into 2.5D multi-layer representations [DJB13], which include surface attributes such as gradient, local curvature, normal, and albedo [Mac15]. A wide variety of RTI acquisition setups exist, ranging from low-cost and transportable kits [CHI16], to different sizes of fixed light domes [CHI16, SSWK13, Ham15]. Recently, some dome solutions have been presented that use both visible and invisible light wavelengths [LG14, Leu15]. Our setup is both low cost, simple to implement, and provides an easy capture procedure. It is not a fixed light setup and allows a more flexible choice of number and position of the light sources, which is very important when dealing with non-Lambertian, shiny materials. Moreover, it is easy to extend the presented pipeline to the multi- or hyper-spectral domain at a much lower cost than multiple-light setups.

Coping with non-ideal lighting. While dome solutions allow for pre-calibration, and thus more reliable data, the main problem in classic RTI approaches with freely movable lights is that they generally rely on some simple and limiting assumptions about lighting and materials. The most common assumption is a Lambertian material under uniform lighting generated by an infinite distance light source (parallel rays). The computation of surface attributes is then prone to errors due to this simplistic model. For this reason, Xie et al. [XDW15] proposed a photometric stereo method that uses local/global mesh deformation approach to solve the non-linear problem of PS with near point lighting. Rather than computing only surface orientation, they rely also on the computation of surface point position, which is a well-known problem in the case of surface discontinuities or self and cast shadows. Moreover, although they take into account the attenuation of luminance due to the distance between surface points and light source, they consider a Lambertian material. Similarly, Ahmad et al. [ASSS14] developed a new procedure to allow dynamic calibration of per-pixel light vectors as an object moves in the field of view. Unfortunately they also rely on Lambertian diffused maxima region. A more controlled, general calibration method that does not depend on Lambertian assumption has been presented recently by Xie et al. [XSJ*15]. They calibrate a LED-based system to take into account near-field lighting and the dependency of surface irradiance on the light distance. They show acquisition of opaque and specular objects. However, their approach is strongly dependent on the model of LED radiance distribution, which is not uniform but isotropic. In a similar manner Ukida et al. [UTSY15] developed an hybrid method (stereo plus photometric stereo) tailored on a illumination model of a linear light source. Conversely, some other advanced methods were proposed by Mecca et al. [MRC15b, MRC15a]; they are based on a more complex mathematical formulation that exploits partial differential irradiance equation ratios to deal with all the mentioned highly non-linear physical factors being involved in the image acquisition process. They model the material as a combination of a diffuse Lambertian and a specular Blinn-Phong components. However the deformation of their results are still large for non-Lambertian objects and the computational time of that more complex algorithm has been tested only with small images. Giachetti et al. [GDR*15], similarly to us, instead of tackling the complex problem of recovering light positions and form factors, directly reduce the inhomogeneity of lighting using a calibration

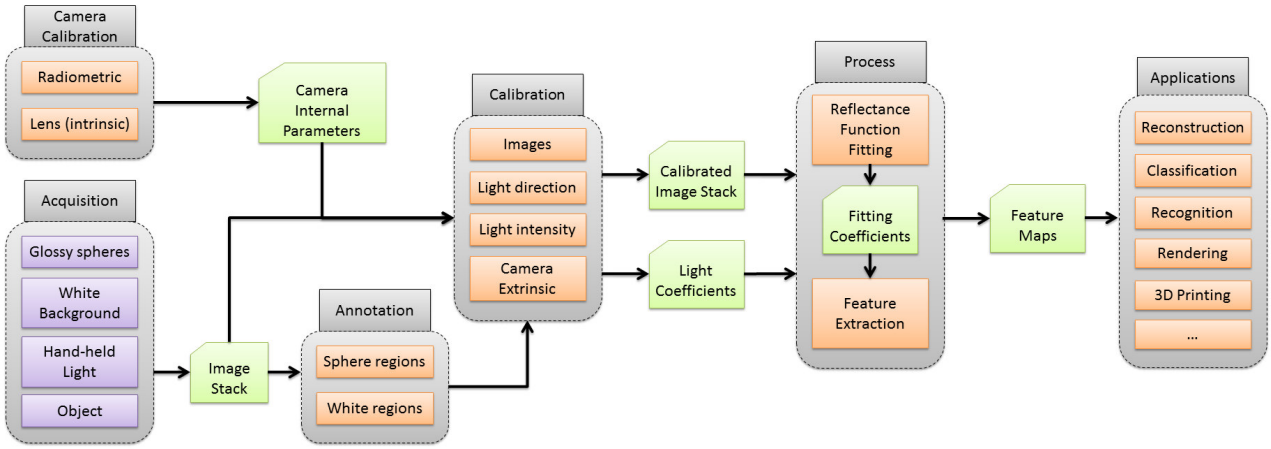


Figure 1: ***Surface acquisition and characterization pipeline.*** *Objects are acquired with a pre-calibrated camera and moving lights. Calibration objects added to the scene to automatically recover light positions and to compensate for non-uniform illumination. The resulting pipeline computes surface features usable for a variety of applications.*

plane with constant reflectivity. Their work, however, is limited to flattening the estimated lighting profile to emulate a purely directional lighting, while we extend the proposed light correction approach to a true per-pixel illumination correction accounting for per-pixel light direction and intensity,

From image stacks to parametric representations, There are various models for fitting the reflectance distribution data obtained with RTI acquisition and this paper focuses on polynomial based interpolation (such as the seminal work of Malzbender et al. [MGW01]). As explained in Drew et al. [DHOMH12], polynomial approximation improves Lambertian Photometric Stereo (PS) by solving a higher order polynomial and, coupled with a dedicated reflectance model, allows for more realistic reproduction of the material's radiance. The biquadratic function in typical polynomial approximations can be solved with linear regression by minimizing the least-squares error or obtaining more robust results by trimming the outliers with least median squares method (LS) [RL05, ZD14, DHOMH12]. The regression outputs a set of 6 coefficients for each pixel in the image dataset, that can be treated as six spatial maps with the same spatial resolution as the original images, but with a reduced resolution in the angular space due to the n image-to-6 approximation [Mac15]. Here we use a trimmed LS [RL05] based on a simple and fast outlier diagnostic strategy. It proves to be efficient even with the considerable amount of black noise and highlights caused by very low diffuse albedo and a high shininess of tested materials. A similar strategy can be applied as well on other fitting functions like Hemispherical Harmonics (HSH) or higher order polynomials [GKPB04, ZD14] that could be easily added to our framework.

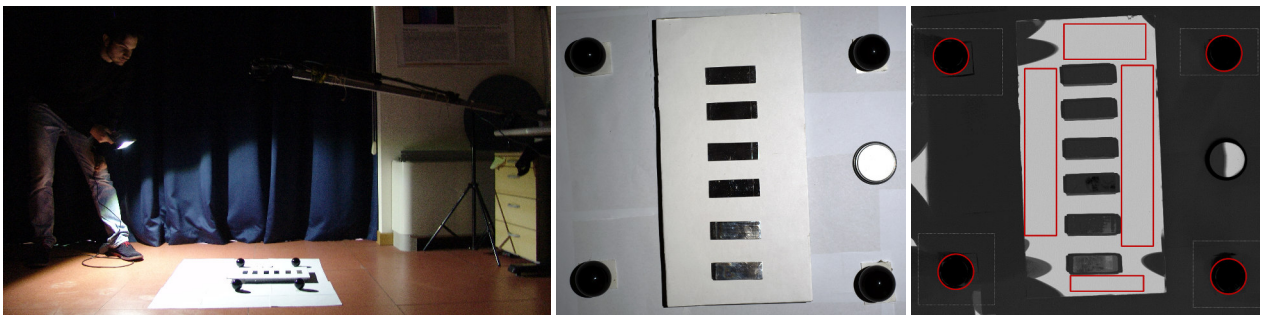
Material classification and descriptors Previous attempts have been made in the direction of material classification based on the fitted coefficients, e.g. HSH [WGSD09], PTM [GDR*15] and PS [TGVG12] coefficients. In cited papers, pointwise reflectance parameters are used to characterize the materials and segment different regions of a planar surface by means of clustering algorithms. This approach has clear limits due to the difficulty in light calibration, the dependency of the reflectance on local geometry, and other complex factors Our idea is to try to overcome these limits using a smart calibration approach.

3. Method

The basic idea behind the presented pipeline is to avoid the necessity of using complex calibrated light setups to have a real control of the reflectance estimation, and to avoid either the large errors in the reflectance parameters estimation due to light inhomogeneity typical of freehand/highlight RTI capture. This is obtained using a capture setup with multiple reflective spheres and a planar white Lambertian target. We propose to pre-calibrate the raw input data before parameter fitting in order to treat each pixel as if we had an ideal constant illumination intensity coming from a known ray direction. We do so by interpolating the light direction over the whole image and using white calibration patches to estimate local illumination intensity at each pixel location. The setup can be easily used in lab environments to acquire a wide range of objects (paintings, bas-reliefs, coins, etc.), but it could be even applied in outdoor acquisitions provided the proper positioning of calibration targets.

An overview of the proposed acquisition and processing pipeline is presented in Fig. 1. Before the acquisition, we have already (and once) calibrated the internal characteristics of the camera, i.e., the radiometric response and the lens parameters. The acquisition setup includes some reflective spheres, a white planar frame as background, an hand-held light (freely moved by the user), and the object being captured. The white planar frame can be substituted with any kind of material with constant reflectance in the frequency band used for illumination.

The capture process outputs an image stack, which is semi-automatically annotated to find the position of the spheres and to segment the white background parts not cluttered by the object. These three inputs (camera internal parameters, image stack, and



(a) Setup (b) Scene (c) Annotation

Figure 2: ***Acquisition and annotation.*** *(a) The acquisition setup composed of the stationary DSLR digital camera, the hand-held light source and the CH targets arranged horizontally on a white background together with the four black reflective spheres. (b) Acquired scene. (c) User interface for the annotation of reflective spheres and frame regions.*

annotation) are used by the algorithm to perform a series of fine calibrations on images, light and camera. After that, each pixel is associated with a calibrated reflectance profile, coupled with calibrated light parameters; those are used to perform a reflectance function fitting in order to extract an objective lower dimensional representation described by a small number of coefficients. These feature maps might be employed in a variety of applications, such as surface reconstruction, object classification and recognition, virtual rendering/re-lighting, 3D printing and so forth. In the following, we describe in more details the various pipeline steps.

3.1. Camera calibration

Reflectance Transformation Imaging uses measurements of irradiance on the camera sensor to extract photometric and geometric information about the viewed scene. For this reason, we calibrate the camera both in terms of radiometric response and geometrical behavior. The radiometric calibration aims at coping with the non-ideal, non-linear response of the camera sensor, and is obtained by photographing a target under different exposures and computing parameters that transform the irradiance signal so that the measured brightness value will be linearly proportional to the amount of photons hitting the sensor [GN04, Yor16]. We also take into account the geometrical behavior of the lens by calibrating the intrinsic parameters of the camera [GoCMM16]. This will be used to remove the radial distortion from the input image stack.

3.2. Acquisition

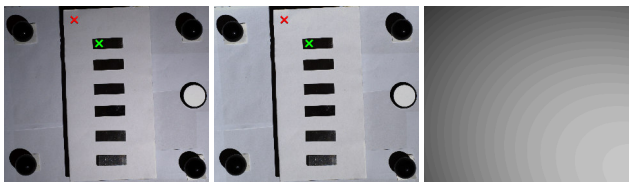
The acquisition setup consists in a camera, a hand-held light, some calibration targets (a white planar frame and four glossy spheres) and the object under study, which is assumed to be approximately flat. In fig. 2(b) we present the scene viewed by the camera, which contains four dark glossy spheres and a white diffuse planar target, on which we positioned six metallic objects being acquired. During the acquisition, for each image the user changes the position of the light, and he tries to uniformly sample the direction across the hemisphere above the scene 2(a). A uniform sampling will assure a well mathematically conditioned light matrix and reduce the computation errors during the reflectance function fitting [KH14]. Apart from that, we don't have any particular constraints on the light form factor (we do not rely on a point light source approximation), nor we force the user to strictly remain at the same distance from the scene to avoid variation of light intensity. Moreover, since we avoid the common parallel rays assumption (or, far light condition), the user can remain close to the object; this will increase the average scene irradiance, decrease the exposure time per image, and then speed-up the entire capturing process. Without loss of generality, in all our experiments the ambient light is completely negligible. If that is not the case, many methods can robustly deal with the presence of ambient light, even if its intensity is many times the illuminance of the directional light component [AP14].

3.3. Annotation and calibration

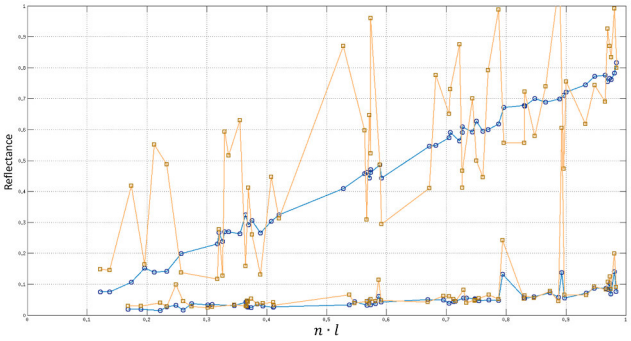
We propose a semi-automatic, simple, and fast calibration pipeline, in order to extract reliable object surface information. It consists in two main steps, i.e., light direction and intensity calibration. Using a custom software, we are able to estimate the light direction from the sphere highlight positions, and we assume that the light direction at each object point is well approximated by a linear interpolation of the estimated values, as shown in [GDR*15]. The user annotates the region of interest (ROI) of each sphere (see fig. 2(c). Then the ball position, the highlight position and the derived light direction associated to the centers of the spheres are estimated. The light direction for each pixel is computed by linear interpolation of those known directions. Now that we have the light direction for each pixel, we need to compute and calibrate the spatially varying light intensity. We exploit the white planar target, since we can make some reliable assumptions about its Bidirectional Reflectance Distribution Function (BRDF), albedo (ρ) and normal (n) values in the viewer-oriented coordinate system. First of all, it is a Lambertian diffuser, so the reflected radiance will be:

$$L_r = \frac{\rho}{\pi} \int_{\omega} L_i(\theta_i, \phi_i) \cos\theta_i d\omega_i \quad (1)$$

Whatever is the nature of our light source, the previous calibration of light directions allows us to consider a per-pixel collimated illumination, and the resulting simplified double-delta representation



(a) Original (b) Calibrated (c) Interpolated γ^{-1}



(d) Original vs. Calibrated Profile

Figure 3: ***Light intensity calibration.*** *In the top row, the original image corresponding to light #6, its calibrated version and the interpolated correction factor. In the bottom, two original (orange) and calibrated (blue) appearance profiles of the pixels indicated by the red and green crosses in the top image. Images sorted by angle with the white calibration plane normal.*

of the source incoming radiance [HS79]:

$$L_i(\theta_i,\phi_i) = E_0\delta(cos\theta_i - cos\theta_0)\,\delta(\phi_i - \phi_0) \tag{2}$$

E_0 is the irradiance produced by a collimated source on a surface oriented orthogonally to the direction (θ_0,ϕ_0), and (θ_0,ϕ_0) is the per-pixel known calibrated light direction. E_0 must be constant and equal to a reference white value ($\rho \leq 1$) for all the pixels in the white diffuse target. However, this is not true due to the spatially varying nature of the light. For this reason, given a pixel (u,v) in the diffuse target, and $\tilde{L}_r(u,v)$ the corresponding real measured radiance, the value of E_0 will be:

$$E_0(u,v) = \frac{\pi\tilde{L}_r(u,v)}{\rho cos\theta_0} \tag{3}$$

Note that we know the value of $cos\theta_0$ since we know the illumination direction computed by interpolation from the directions recovered from sphere highlights, and we can recover the true white planar target normal, which is approximately collinear with the viewing direction, i.e., $n = (0,0,1)^T$, from the known spherical target sizes and positions in the image. At this point, we can easily find the new calibrated pixel value $L_r(u,v)$ for the white target region by computing $\tilde{L}_r(u,v)\gamma(u,v)$, where $\gamma(u,v) = E_0(u,v)^{-1}$. Using this value instead of the original one for each of the acquired images produces a calibrated appearance profile that is independent of effects such as near-light and non-uniformity of lamp illumination.

Now that we are able to retrieve the calibrated pixel value for the white region, we need to find a procedure to apply the same process to the rest of the image, and in particular to the objects of interest. Given a sparse sampling of light intensity, we assume that its variation in the image space is not a high frequency signal, but it can be well approximated by smooth function. Moreover, we assume that objects are thin and flat, thus approximately at a small distance from the plane defined by the white diffuse target. This makes a cubic, plane-based fitted signal still valid for the object points. From a practical point of view, we need to segment and consider the set of pixels in the white region that are not in shadow, which are used to compute the known values of γ. Then, in a completely automatic way, the algorithm computes the values of γ for the unknown pixels by fitting a cubic polynomial. The final pixel values are obtained by pixel-wise multiplication of the $\gamma(u,v)$ field by the measured image $\tilde{L}_r(u,v)$. Fig. 3 shows the effect of light intensity calibration on one of the original image in the input stack. We obtained the calibrated image (Fig. 3(b)) by multiplying the original image (Fig. 3(a) with the interpolated γ function rescaled to the unit interval to make it more readable (Fig. 3(c)). The effect of light intensity correction is pretty evident across all the image, and, in particular, it is very clear in the upper left corner. There the original image is very dark mostly due to the combination of spatially-varying light (spot effect), and the distance from the light source (near- and point-light effect). Since for each acquisition the light is in a different position with respect to the scene, deviations across images with respect to the constant illumination assumption can be dramatic.

Since this calibration takes into account radiometric nature of both the image-forming system and the known targets, the calibration of light intensity compensates the spatially varying behavior keeping reflectance values photometrically consistent across all the input image stack. The result of the entire calibration procedure is a set of corrected reflectance appearance profiles (calibrated image stack), and a set of light coefficients to retrieve the per-image and per-pixel light directions. The effect of calibration is evident in Fig. 3(d), where we present the appearance profile data of the two pixels marked with the red (white target) and green (sample) crosses in Fig. 3(a) and Fig. 3(b). We can see that uncalibrated appearance profiles look extremely jaggy due to unwanted lighting effects, and would lead to noticeable difficulties and errors in surface and reflectance parameters estimation.

Figure 4: ***Enhancement****. We compare the original picture (left) with two different kinds of enhanced visualization of the captured model, i.e., unsharp masking applied to the normal field (center) and to the full set of PTM coefficients (right). Note how enhancement makes severely degraded details readable.*

3.4. Surface characterization

The calibrated appearance profiles can be used for extracting stable parameters characterizing the object surface. Quantitatively speaking, the profile signal is a sum of several contributions, ranging from the diffuse behavior of the material to its specular component. In this paper, instead of modeling the high frequency modeling of the material Bidirectional Reflectance Distribution Function, we explore the possibility of extracting meaningful information from the low frequency part, which consists in all the Lambertian and non-Lambertian diffuse components [IA14].As discussed in Sec. 4, this part of the signal is reliable enough even in the case of highly specular metallic objects. The low frequency assumption allows us to approximate per-pixel reflectance profiles by low-degree polynomial interpolation [MGW01]. We employ the polynomial representation presented by Drew et al. [DHOMH12], which uses 6 coefficients (a,b,c,d,e,f) to relate the light direction (l_x,l_y,l_z) with the measured, calibrated reflectance L_r:

$$L_r = [a,b,c,d,e,f]\left[l_x,l_y,l_z,l_x^2,l_xl_y,1\right]^T \tag{4}$$

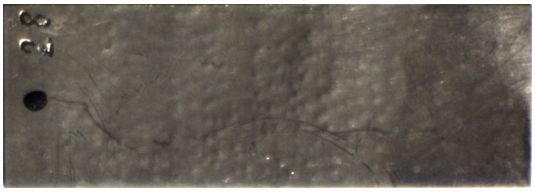
The first three coefficients (a,b,c) model the pure Lambertian part of the reflectance, the coefficient f models the ambient light term, while the two coefficients (d,e) take into account other diffuse and low-frequency non-diffuse surface behavior (e.g., non-Lambertian diffuse signals or interreflections).

Given a profile, in order to remove outliers due to shadows and high-frequency reflections, we apply Trimmed Least Squares [RL05]) to fit the data with the equation 4. For each pixel, we estimate the 6-dimensional vector of polynomial coefficients, which we consider our lower dimensional representation of the appearance profile. This way to express per-pixel data is very useful for different kind of applications, e.g., it is possible to directly use the coefficients as a $6D$ texture, and to compute descriptors from it for classification and recognition purposes [BCDG13, GDR*15]. Here we focus on the reconstruction of normal and albedo of the surface from the coefficients. The first three coefficients (a,b,c) represents the albedo-scaled normal vector, so its magnitude is the albedo, while the direction is the normal. Computing the normals with a polynomial interpolation rather than the classic photometric stereo 3-dimensional model takes into account not only the Lambertian component of the material but a more wide diffuse non-Lambertian behavior, resulting in a more relaxed set of assumptions [IA14]. Moreover, having a precise and reliable estimation of those surface attributes is extremely important to obtain information about surface behavior (e.g., roughness [Kay15]), and, by integrating normals, to compute a 3D model as well [BB15]. We will show in the section 4 how the proposed, simple calibration pipeline provides more stable quantitative results with respect to standard free-form RTI approaches, by producing repeatable and reliable outcomes that are comparable with other common and established measuring systems (i.e., the optical microprofilometer). Moreover, by suitably processing the coefficients, it is possible to render the objects in order to enhance particular aspects of them, making them more readable. An example is presented in Fig. 4, where a severely damaged coin is made more readable by enhancing local variations. Note how, by unsharp masking the full PTM coefficients, we obtain a more readable result than just unsharp masking the normal.

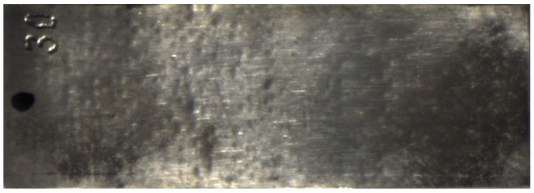
(a) It. Bronzital 10c (b) It. Copper 10c (c) Roman Quad.

(d) Silver sample #15

(e) Silver sample #22



(f) Silver sample #28



(g) Silver sample #30

Figure 5: ***Samples.*** *Photos of the metallic samples used to test the proposed approach: (a) an Italian Bronzital 10c coin; (b) a severely degraded Italian copper 10c coin; (c) a bronze roman coin; one polished silver with a partial coating treatment(d); three chiseled silver plates, i.e., #22(e), #28(f) and #30(g).*

4. Results

The pipeline described in this paper has been implemented using widely available devices, including a LED lamp and a DSLR camera. We present in this work the results obtained in the reconstruction and analysis of some metallic samples, which, due to their

sometimes very high reflectivity and low diffusivity, present a difficult test case for RTI approaches.

4.1. Samples

We tested our system on a variety of samples and cultural objects. We present in this paper results obtained on laboratory-created metallic plates, as well as on ancient and antique coins (see Fig. 5).

The plates, custom made to study metal and metal treatment behavior, are made of 95% silver and 5% copper. They have a thickness of 1mm and a size of 7cm by 2.5cm. Three of the tested samples (sample 22, 28, 30) were chiseled, in order to emulate the appearance of cultural objects, while the fourth one (sample 15) is polished and partially coated with nitrocellulose and wax. The chiseled samples are used in this paper to test the geometric accuracy and repeatability of our pipeline, while the polished one is used to test the repeatability of surface characterization.

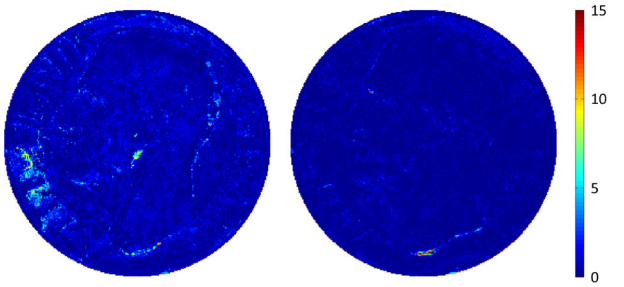
The coins are a small bronze roman coin (*quadrans*) dated 9 BC. and damaged by scratches, and two 10 cent Italian coins. One exemplar, dated 1931 is made of copper and is severely degraded, while the second exemplar, dated 1939, is made of a special alloy with nickel called Bronzital, which has been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

4.2. Camera calibration

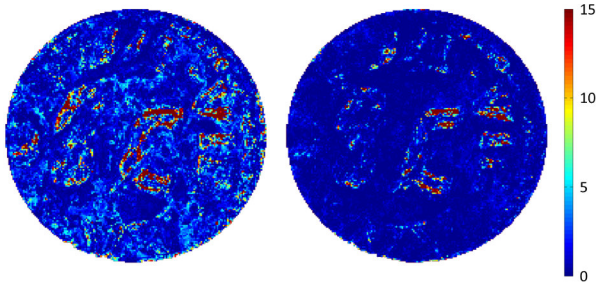
The digital camera used for the experiment is a DSLR Nikon D810 with CMOS sensor (36x24mm, spatial resolution of 36MP for the full format image area). The lens attached was a full resolution AF-S FX Nikkor 50mm f/1.8G. The sensor of the digital camera was checked for linearity, by taking images with different exposures, from very dark to very bright and then plotting the brightness of the images as a function of exposure. The camera was geometrically calibrated by computing intrinsic parameters and undistorting the images before feeding them to the RTI pipeline. We employed a model with two radial and two tangential distortion coefficients. We used the GML Camera Calibration Toolbox [Vel16], but any other calibration tool can be used as well to perform this task.

4.3. Acquisition setup and protocol

The acquisition scene was set-up horizontally and the digital camera was placed on top at a distance of approximately 95cm, on a perpendicular direction to the image plane, as shown in Fig.2(a). The Ground Sampling Distance when using the 50 mm lens was 0.009 cm/pixel. The metallic samples were arranged on a white paper frame. Four reflective black spheres were included so as to enable light direction estimation from the freely moved hand-held light source. A Spectralon target was included in the scene as a perfect white diffuser. The light source used for the experiment was a white LED (color temperature 6500K) for visible (VIS) photography, with very limited emission in the infrared (IR). The digital camera was remote controlled. The acquisition configuration was set as follows: ISO 32, aperture f8, shutter speed 0.8s, custom white balance and manual focus. The samples were acquired in two sessions. During each acquisition, a virtual light dome was created by manually moving the light source around the samples, from 12 azimuth locations with 4 corresponding elevation points and a final acquisition from the top, giving a total of 49 images which can translate into 49 light directions. All images were shot in RAW format.

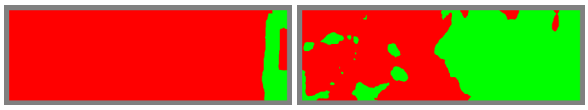


(a) Italian Bronzital 10c coin

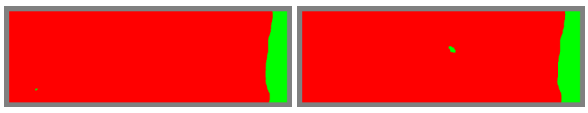


(b) Roman Quadrans

Figure 6: ***Normal deviation in two different acquisitions.*** *For each coin, the left image maps the angular deviation of normals reconstructed from two different acquisitions without light calibration, while the right image presents the result obtained when light direction and intensity are corrected per pixel.*



(a) Without calibration



(b) With calibration

Figure 7: ***Unsupervised classification****. Two class k-means clustering applied to two different acquisitions of the polished sample #15 (fig. 5(d)). The coated area is in red while the uncoated area is in green. Without calibration we have* 20% *of classification similarity, while we obtain a value of* 99.5% *by using the calibrated images. This shows the drastically increased level of repeatability of the proposed pipeline with respect to classic free-form RTI.*

4.4. Repeatability and characterization

In order to test for repeatability, acquisitions of the same samples were repeated multiple times, and the PTM coefficients were reconstructed using a robust approach with trimming with and without light intensity and direction calibration before parameter fitting.

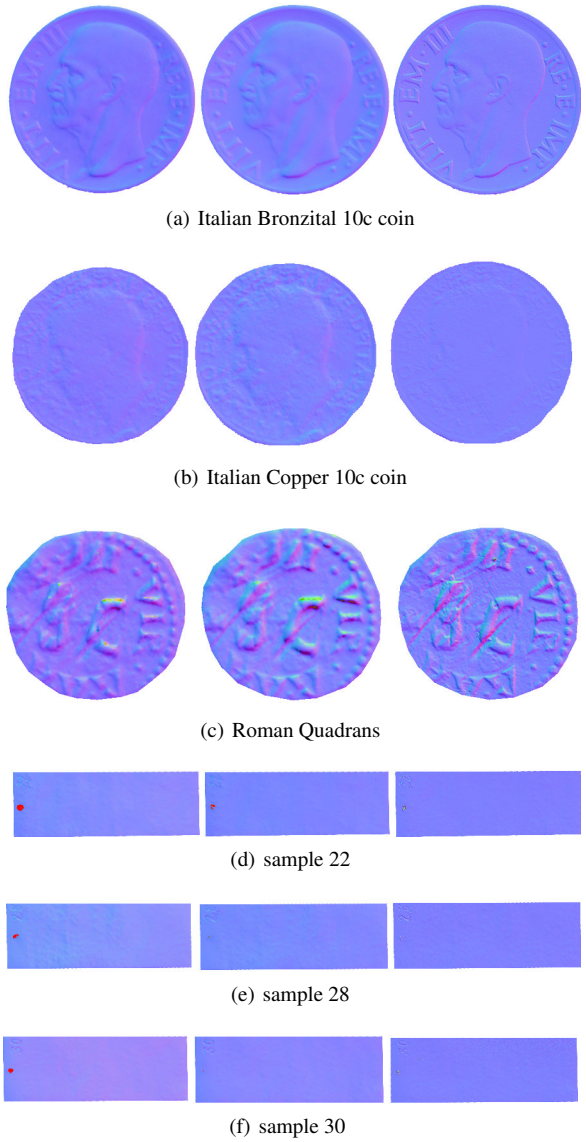
As expected, light calibration sensibly improves acquisition, making it more predictable. Fig. 6 shows, for the Roman and Italian coin samples, the deviation in the reconstructed normals acquired twice using the same number of lights and a similar illumination pattern, controlled free hand. The median angular deviation of normals decreases, when using calibration, from 1.86 to 1.17 degrees for the Bronzital 10c coin and from 2.69 to 1.41 degrees for the Roman Quadrans. Much larger variations occur on the metallic plates, which have a larger extent and, therefore, suffer more of light inhomogeneity.

The repeatability of the proposed approach is particularly important to obtain a stable surface characterization. As a representative example of material classification, we have performed two RTI acquisitions of a polished silver sample partially covered by a coating, see Fig. 5(d), and applied unsupervised classification to segment regions with or without coating. For each RTI sample, we compute a 7-dimensional descriptor of a 30 pixels neighborhood. The descriptor is the average albedo value, to account for material color, plus the 6 standard deviations of the polynomial coefficients, to account for the roughness of the sample surface.

Unsupervised classification is achieved by performing two-class k-means clustering. We measure the similarity of the classification outcomes obtained from the two different acquisitions, without and with light calibration, see Fig. 7. The only difference between the two acquisitions of the same sample is the different lighting pattern caused by the free-form approach. The coated area is in red while the uncoated area is in green. In the absence of light calibration, the clustering outcome is unstable, as it has only a 20% overlap, while, by performing light calibration, we improve it up to 99.5% of pixels that have been assigned to the same class, showing that with our approach free-form RTI can be used for characterization.

4.5. Comparison with microprofilometer

In order to test the accuracy of shape capturing, we compared our reconstruction with measures taken with an optical microprofilometer based on conoscopic holography [SP85], which is capable to take reliable and stable profilometric measurements down to the scale of micron on different kinds of materials, reflective or diffusive. The microprofilometer is based on an Optimet conoscopic probe mounted on sub-micrometric linear stages in order to scan a region up to $30x30cm^2$ in one session. Different acquisition setups are used to adapt the working range of the system to the deformation of the various objects, with the surface data collected at different spatial grid. Silver samples, which are highly reflective and "flat" surfaces (submillimetric deformation), were measured with a transversal resolution (XY grid) of 50 microns with a repeatability (Z heights) of 0.6 microns. Coins, which are less glossy surfaces with deformation in the order of millimeters, were measured with a transversal resolution (XY grid) of 50 microns with a repeatability (Z heights) of 0.1 microns.



(a) Italian Bronzital 10c coin

(b) Italian Copper 10c coin

(c) Roman Quadrans

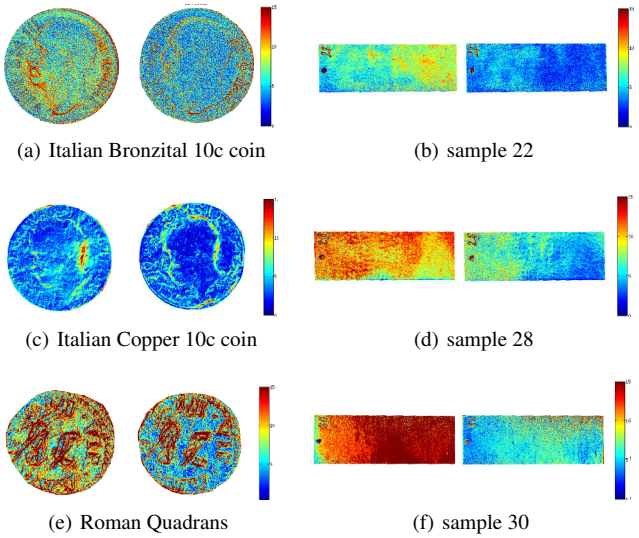
(d) sample 22

(e) sample 28

(f) sample 30

Figure 8: ***Registered normal maps.*** *Color-coded normal maps obtained with the RTI approach compared with microprofilometric reference on the three coins surfaces and the three silver samples. From left to right: uncalibrated RTI, calibrated RTI, ground truth from microprofilometry.*

For our comparison, we measured differences between normal maps estimated from the RTI procedure with those derived by the microprofilometric map. Since normals are very sensitive to high frequency details, even the smallest errors are magnified, and normals are, therefore, a very good indicator. It should be noted that the surface layers captured by different instrument may be slightly different, and the direct comparison should be considered with care due to the different levels of detail of the instruments and their different sensitivity with respect to materials. However, a qualitative and quantitative comparison of the normals can give a reasonable



(a) Italian Bronzital 10c coin (b) sample 22

(c) Italian Copper 10c coin (d) sample 28

(e) Roman Quadrans (f) sample 30

Figure 9: ***Angular errors***. *Color-coded angular errors (degree) of RTI estimated normals wrt ground truth from microprofilometry. Left: non-calibrated results. Right: results with light calibration.*

idea of the quality of our results. To obtain a normal map comparable with ours from microprofilometer depth maps, we registered the depth over the estimated RTI normal maps with a similarity transform optimized to match the correspondence of manually selected points (12 landmarks). This initial registration was refined by locally optimizing mutual information in image space. In order to reduce inaccuracies due to human intervention, we performed the alignment procedure several times for each sample and kept the best result. Very small differences among trials were measured.

Fig. 8 shows the normal maps of the objects in Fig. 5 mapped in the same reference system: the left images are the RTI normals obtained with original images, the center images are those obtained with calibrated photos, while the right column shows the normals estimated from microprofilometry. It is possible to see the good quality of the reconstructed maps, even in the case of the severely damaged and dark copper coin, where details are not easily visible in the images. Moreover, very good results are also obtained on highly specular objects, such as the silver samples and the Bronzital coin. From the quantitative point of view, light calibration sensibly improves reconstruction quality, as reported in Table 1 and Fig. 9, which illustrate the angular differences with respect to the microprofilometer obtained with and without calibration. The calibration procedure reduces the median errors on average of 27 percent.

5. Conclusions

We presented a practical acquisition and processing pipeline based on free-form Reflectance Transformation Imaging to characterize the surface structure of cultural heritage objects. The basic idea behind the presented pipeline is to simplify capture and increase its flexibility by avoiding the necessity of using complex calibrated light setups, while ensuring, at the same time, avoiding the large errors in the parameters estimation due to light inhomogeneity typical of freehand/highlight RTI capture. This is obtained by using a capture setup with multiple reflective spheres and white Lambertian surfaces from which to automatically recover light positions and to compensate for non-uniform illumination. As demonstrated in Section 4, the results are promising even for very difficult surfaces such as the metals tested in this work, since the measured quantitative parameters are improved with respect to the current non-corrected capture setups, the geometric surface characteristics are in-line with those obtained with a micro-profilometer, and simple features derived from measurement lead to good discrimination of surface conditions.

Table 1: *Median angular distances of the RTI estimated normals from the reference microprofilometer normals. The calibration procedure reduces the errors on average of 27 percent.*

	Median angular distance (deg.)	
	Non-calibrated	Calibrated
Bronzital 10c	6.69	4.51
Copper 10c	3.89	3.03
Quadrans	9.77	6.17
Sample #22	7.07	3.53
Sample #28	10.45	5.44
Sample #30	14.60	5.56

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

[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. 2

[AP14] ANGELOPOULOU M. E., PETROU M.: Evaluating the effect of diffuse light on photometric stereo reconstruction. *Machine vision and applications 25*, 1 (2014), 199–210. 4

[ASSS14] AHMAD J., SUN J., SMITH L., SMITH M.: An improved photometric stereo through distance estimation and light vector optimization from diffused maxima region. *Pattern Recognition Letters 50* (2014), 15–22. 2

[BB15] BÄHR M., BREUSS M.: An improved eikonal method for surface normal integration. In *Pattern Recognition*. Springer, 2015, pp. 274–284. 6

[BCDG13] BROGNARA C., CORSINI M., DELLEPIANE M., GIACHETTI A.: Edge detection on polynomial texture maps. In *Image Analysis and Processing–ICIAP 2013*. Springer, 2013, pp. 482–491. 6

[CHI16] CHI: Cultural heritage imaging, 2016. URL: http://culturalheritageimaging.org. 2

[DHOMH12] DREW M. S., HEL-OR Y., MALZBENDER T., HAJARI N.: Robust estimation of surface properties and interpolation of shadow/specularity components. *Image and Vision Computing 30*, 4 (2012), 317–331. 3, 6

[DJB13] DUFFY S. M., JONES D., BACKHOUSE P.: *Multi-light imaging for heritage applications*. 2013. 2

[DRS10] DORSEY J., RUSHMEIER H., SILLION F.: *Digital modeling of material appearance*. Morgan Kaufmann, 2010. 2

[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. 2, 3, 4, 6

[GKPB04] GAUTRON P., KRIVANEK J., PATTANAIK S. N., BOUATOUCH K.: A novel hemispherical basis for accurate and efficient rendering. *Rendering Techniques 2004* (2004), 321–330. 3

[GN04] GROSSBERG M. D., NAYAR S. K.: Modeling the space of camera response functions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 26*, 10 (2004), 1272–1282. 4

[GoCMM16] GRAPHICS, OF COMPUTATIONAL MATHEMATICS M. L. F., MOSCOW C.: Gml c++ camera calibration toolbox, 2016. URL: http://graphics.cs.msu.ru/en/node/909. 4

[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. 2

[HS79] HORN B. K., SJOBERG R. W.: Calculating the reflectance map. *Applied optics 18*, 11 (1979), 1770–1779. 5

[IA14] IKEHATA S., AIZAWA K.: Photometric stereo using constrained bivariate regression for general isotropic surfaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2014), pp. 2179–2186. 6

[JMS*14] JOHNSON C. R., MESSIER P., SETHARES W. A., KLEIN A. G., BROWN C., DO A. H., KLAUSMEYER P., ABRY P., JAFFARD S., WENDT H., ET AL.: Pursuing automated classification of historic photographic papers from raking light images. *Journal of the American Institute for Conservation 53*, 3 (2014), 159–170. 1

[Kay15] KAYA B.: Surface roughness inspection in milling operations with photometric stereo and pnn. *The International Journal of Advanced Manufacturing Technology 81*, 5-8 (2015), 1215–1222. 6

[KH14] KLAUDINY M., HILTON A.: Error analysis of photometric stereo with colour lights. *Pattern Recognition Letters 48* (2014), 81–92. 4

[KN06] KOPPAL S. J., NARASIMHAN S. G.: Clustering appearance for scene analysis. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on* (2006), vol. 2, IEEE, pp. 1323–1330. 2

[Leu15] LEUVEN K.: Multispectral microdome, 2015. URL: https://portablelightdome.wordpress.com/2015/04/29/rich-presents-the-new-multispectral-microdome/. 2

[LG14] LIU C., GU J.: Discriminative illumination: Per-pixel classification of raw materials based on optimal projections of spectral brdf. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 36*, 1 (2014), 86–98. 2

[Mac15] MACDONALD L. W.: *Realistic visualisation of cultural heritage objects*. PhD thesis, UCL (University College London), 2015. 2, 3

[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, 3, 6

[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

[MRC15a] MECCA R., RODOLÀ E., CREMERS D.: Analysis of surface parametrizations for modern photometric stereo modeling. In *The International Conference on Quality Control by Artificial Vision 2015* (2015), International Society for Optics and Photonics, pp. 95341B–95341B. 2

[MRC15b] MECCA R., RODOLÀ E., CREMERS D.: Realistic photometric stereo using partial differential irradiance equation ratios. *Computers & Graphics 51* (2015), 8–16. 2

[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

[PCC*10] PALMA G., CORSINI M., CIGNONI P., SCOPIGNO R., MUDGE M.: Dynamic shading enhancement for reflectance transformation imaging. *Journal on Computing and Cultural Heritage (JOCCH) 3*, 2 (2010), 6. 1

[RL05] ROUSSEEUW P. J., LEROY A. M.: *Robust regression and outlier detection*, vol. 589. John Wiley & Sons, 2005. 3, 6

[SP85] SIRAT G., PSALTIS D.: Conoscopic holography. *Optics letters 10*, 1 (1985), 4–6. 8

[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. 2

[TGVG12] TINGDAHL D., GODAU C., VAN GOOL L.: Base materials for photometric stereo. In *Computer Vision–ECCV 2012. Workshops and Demonstrations* (2012), Springer, pp. 350–359. 3

[UTSY15] UKIDA H., TANIMOTO Y., SANO T., YAMAMOTO H.: 3d shape and color estimation using linear light sources and cameras. In *Imaging Systems and Techniques (IST), 2015 IEEE International Conference on* (2015), IEEE, pp. 1–5. 2

[Vel16] VELIZHEV A.: Gml c++ camera calibration toolbox, 2016. URL: http://graphics.cs.msu.ru/en/node/909. 7

[WGSD09] WANG O., GUNAWARDANE P., SCHER S., DAVIS J.: Material classification using brdf slices. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* (2009), IEEE, pp. 2805–2811. 3

[WK15] WEINMANN M., KLEIN R.: Advances in geometry and reflectance acquisition (course notes). In *SIGGRAPH Asia 2015 Courses* (2015), ACM, p. 1. 2

[Woo78] WOODHAM R. J.: Photometric stereo. 1

[XDW15] XIE W., DAI C., WANG C. C.: Photometric stereo with near point lighting: A solution by mesh deformation. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on* (2015), IEEE, pp. 4585–4593. 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. 2

[Yor16] YORK C. V. L. C. U. N.: Rascal - radiometric self calibration, 2016. URL: http://www1.cs.columbia.edu/CAVE/software/rascal/rrhome.php. 4

[ZD14] ZHANG M., DREW M. S.: Efficient robust image interpolation and surface properties using polynomial texture mapping. *EURASIP Journal on Image and Video Processing 2014*, 1 (2014), 1–19. 3