Objective and Subjective Evaluation of Virtual Relighting from Reflectance Transformation Imaging Data

R. Pintus[1], T. Dulecha[2], A. Jaspe[1], A. Giachetti[2], I. Ciortan[2], E. Gobbetti[1]

[1] CRS4, Italy
[2] University of Verona, Italy

Abstract

Reflectance Transformation Imaging (RTI) is widely used to produce relightable models from multi-light image collections. These models are used for a variety of tasks in the Cultural Heritage field. In this work, we carry out an objective and subjective evaluation of RTI data visualization. We start from the acquisition of a series of objects with different geometry and appearance characteristics using a common dome-based configuration. We then transform the acquired data into relightable representations using different approaches: PTM, HSH, and RBF. We then perform an objective error estimation by comparing ground truth images with relighted ones in a leave-one-out framework using PSNR and SSIM error metrics. Moreover, we carry out a subjective investigation through perceptual experiments involving end users with a variety of backgrounds. Objective and subjective tests are shown to behave consistently, and significant differences are found between the various methods. While the proposed analysis has been performed on three common and state-of-the-art RTI visualization methods, our approach is general enough to be extended and applied in the future to new developed multi-light processing pipelines and rendering solutions, to assess their numerical precision and accuracy, and their perceptual visual quality.

CCS Concepts

*●**Human-centered computing** → Empirical studies in visualization; ●**Computing methodologies** → Visual inspection; Image representations;*

1. Introduction

The interactive inspection of digital representations of artworks is of fundamental importance in the daily activity of Cultural Heritage (CH) scholars, as it can reveal geometric cues or precious information about the materials used to create an object, supporting the definition of conservation and preservation strategies. Moreover, digital representations are also increasingly used to present cultural objects to experts or to a wider public, in order to replace, augment, or complement the inspection of real objects.

Among all methods to acquire, process and visualize virtual replicas, Reflectance Transformation Imaging (RTI) is one of the most popular in CH. This is due to the low acquisition hardware cost and to the simple and fast capture and processing pipeline. In addition, RTI is a technique that naturally supports interactive relighting, a type of visualization very appropriate to inspect fine surface details and resembling the classical physical inspection raking light sources to reveal surface detail of actual objects under study. For this reason, RTI visualization is broadly used for relighting images of a wide range of items, e.g., coins [MVSL05, KK13], rock art [Duf10, UW13], paintings [MBW*14], bas-relief [PCC*10], and many more CH items [AIK13].

In the most common RTI capture, process and visualization pipeline, a series of images is taken under the same fixed view point but with different lighting conditions. The acquired image stack describes for each pixel the so called appearance profile, which is a sequence of measurements of per-point reflectance response. This data is analyzed and visualized by fitting a model that permits to go from the sparse measured samples to a continuous description of the reflected light. The goal of these models is to find the optimal trade-off between the amount of input data (sparse light sampling vs high-density gonioreflectometer acquisition), the complexity of reflectance model (computational effort vs model accuracy), the amount of data compression (e.g., for remote/web-based transfer and visualization), rendering interactivity (real-time relighting vs off-line rendering). The final result is a set of parameters that provide a virtual relighting environment in which the user can move a virtual light source in positions different from those sampled. The quality of the RTI visualization depends on how expressive, realistic and artifact-free the rendering is while the user inspects the artwork by moving that virtual light source.

In this work, we carry out an objective and subjective evaluation of RTI data visualization. Using a series of objects with different geometry and appearance characteristics, we acquire them

using a common dome-based configuration. We then transform the acquired data into relightable representations, using different approaches. The resulting relightings, from a variety of light directions, are then evaluated using both objective and subjective tests, aimed at comparing each technique with ground-truth images, as well as to compare each relighting technique with respect to the others. While we present results for selected widely used and state-of-the-art RTI processing and visualization methods, the proposed investigation rationale for RTI quality assessment goes beyond that choice, and can be applied in the future as new solutions will be available. To the best of our knowledge, this is a first attempt to perform an objective/subjective evaluation that systematically investigates the quality of different RTI techniques, applied to the CH field, with both error/perceptual metrics and a web-based visual test performed by different user groups, including experts and non-expert in the visual inspection of CH objects.

2. Related work

The acquisition, processing, analysis and visualization of RTI data is a well known and broadly studied research field. A complete review of this topic is out-of-the-scope of this work, and for a reliable coverage of this research area we refer the reader to recent renowned surveys [Sze10, AG15, DRS10, PPY*16]. Here, after a general description of the acquisition setups, we focus on the literature on multi-light data processing for visualization, which serves as a state-of-the-art background for the proposed evaluation of relighting approaches based on RTI data.

The seminal insight behind RTI [MGW01] was to replace classical multi-light processing based on Photometric Stereo [Woo80] with a more flexible interpolation-based approach that transforms the high amount of data in RTI image stacks into a compact 2.5D multi-layered representation directly usable by applications (e.g., CH object inspection [Duf13]). Those layers might include geometric-based attributes (e.g., normal maps, gradient vector field, local curvature), or appearance based quantities (e.g., albedo maps, per-pixel specular component) [Mac15]. For some applications, this type of approach proved to be a good trade-off compared to more complex 4D Bidirectional Reflectance Distribution Function (BRDF) or Bidirectional Texture Function (BTF) captures. For this reason, the ability of this technique to easily convey meaningful visual information of acquired objects fosters the development of a wide variety of acquisition setups. They range from low-cost, transportable kits [CHI17] to more complex light domes [CHI17, SSWK13, Ham15]. Apart from the way they organize light distribution (i.e., movable or fixed) some solutions increase the number of captured light frequencies in a multi- or hyper-spectral setup [LG14, Leu15].

RTI processing is typically a data reduction procedure, where each N-dimensional per-pixel data (appearance profile) is converted in a parametric representation, where the number of parameters is low compared to the N acquired images. Given an analytical model, RTI processing consists in a data fitting problem. The seminal work [MGW01] uses a bi-quadratic polynomial with coefficients found by linear regression. After that, a lot of different solutions have been proposed in the literature, whose models include spherical and hemispherical harmonics [MMC*08, GKPB04],

eigen hemispherical harmonics [LLW12], bi-polynomial functions [STMI14], discrete modal decomposition [PLGF*15], and bivariate Bernstein polynomials [IA14]. They are all prone to errors due to the fitting of a matte model with an input data that can contain a relevant number of outliers (e.g., produced by shadows and highlights). For this reason, recent methods exploit Robust regression [RL05] to extract outlier free input data to increase the reliability and repeatability in the computation of diffuse model parameters [PGPG17, ZD14]. In addition, some methods obtain alsa the high frequency material behavior by fitting the data with other types of analytical models, e.g., Radial Basis Functions (RBFs). Drew et al. [DHOMH12] use RBFs to fit the difference between the original and the computed matte signals, while Giachetti et al. [GCD*17] use RBFs to produce relighted images by directly processing the original raw RTI stack. The output is used for several applications ranging from visualization [Mac15] to material classification [WGSD09, GDR*15, TGVG12], and feature extraction (e.g., edge detection [BCDG13, Pan16]) and enhancement [CHI17, RTF*04, FAR07].

Previous studies have already shown the interest of RTI analysis for conservation, and several tests have been made to measure repeatability and appropriateness of RTI for conservation tasks [Pay13]. In the application scenario of RTI-based reconstruction and rendering, some state-of-the-art contributions provide a few comparative results. Shi et al. [SWM*16] propose the *DiLiGenT* benchmark and detailed evaluation results that compare various Photometric Stereo approaches with a wide range of optical properties; however, that work assesses the quality of geometric attribute computation (i.e., normal map), leaving the visualization aspect as a complementary unexplored aspect of multi-light processing. The *DiLiGenT* benchmark has also been used to evaluate PTM fitting quality [PGPG17]. Conversely, other works validate specific approaches for RTI data visualization, which improve the output quality of fitting techniques [GCD*17], compression algorithms [PCS18], or enhanced non-photorealistic renderings [BSMG05]. However, although useful, they mostly present sparse and limited comparative results, without a wide coverage of relighting conditions (e.g., light direction range) and material behavior. Moreover, none of them couples an objective set of error metrics together with a human-based visual evaluation provided by experts in the cultural heritage field. For those reasons, in this work we focus our attention on the virtual relighting application, and we proposed an extensive evaluation, both objective and subjective, that aims at assessing which techniques are more suitable to process RTI raw image stacks, and to convey the proper surface appearance visualization for computer-aided object inspection.

3. Overview

To perform the proposed objective and subjective evaluation, we selected four objects with different types of shape and appearance (Sec. 4). We acquired their RTI image stacks with the same fixed light dome (Sec. 4), to avoid biases due to variations in the acquisition conditions. We then process raw data with three different techniques capable to produce relightable representations (Sec. 6). The first two, i.e., classical Polynomial Texture Map (PTM) [MGW01] and HSH relighting [MMC*08, GKPB04], are standard approaches.

Figure 1: *Objects. Original images of the objects chosen for relighting test. First Row: two metallic coins (*Coin1,Coin2*). Second Row: an ancient gold lamina (*Lamina*), and an imprint of a shell (*Shell*).*

The third one, i.e., Radial Basis Functions (RBFs) [GCD*17] has been recently introduced to the RTI domain, and was chosen since it differentiates from the others in the way it computes model parameters by favoring local fitting in the light direction space rather than a global regression. The relighted images produced by the three computed models are compared through an objective and a subjective approach. For the objective comparison, we compare error measurements (PSNR) and perceptual metrics (SSIM) through a leave-one-out evaluation strategy (Sec. 7). We provide also a subjective evaluation by submitting a visual questionnaire to two groups of volunteers: experts and non-expert in CH visual assessment (Sec. 8). The respondents evaluated the effectiveness of each selected technique by providing preferences, as well as estimations of the capability to reproduce ground truth images.

4. Objects

In order to evaluate the quality of the chosen relighting techniques, we consider RTI acquisitions of four objects made of materials that are relevant for the Cultural Heritage domain: two coins made of one and two different metal alloys, an ancient gold lamina, and the imprint of a shell (see Figure 1). In general, we have chosen those objects because they all exhibit different micro/meso-geometry, and different optical material behaviors (both specular and diffuse response). The first coin (*Coin1*) is a quarter US dollar. It has a diameter of 24.26*mm* and a thickness of 1.75*mm*, and it is composed of silver, copper and nickel. The second coin (*Coin2*) is a Polish coin made up of metal, Bi-Metallic Copper-nickel center in Aluminum-bronze ring, and it has a diameter of 21.55*mm* and a thickness of 1.97*mm*. We chose those coins because of their high specular characteristics which allows us to examine the performance of the aforementioned algorithms for the case of specular materials. The other metallic object is a piece of a Phoenician gold lamina, named (*Lamina del Sulcis*). It is a fragment of an inscription engraved on thin gold leaf (a thin plaque or panel intended to be affixed to some other surface), that was found in the earliest stratum of the Sulcis tophet or infant cremation cemetery (West Sardinia). Only 1.4*cmx*1.5*cmx*0.05*cm* in size, the lamina seems to have been

once attached to an iron object, which has partially damaged the surface. The text has been dated to the $8^{th} - 7^{th}$ centuries BCE on the basis of its paleographic features [Bar65]. This makes the object extremely important in Mediterranean archeology [Dix13]. In addition to specular materials, we also examine RTI on matte object, an imprint of shell on a plaster (*Shell*). The object has a very cooperative material, but a difficult geometry, with long and thin curved concave and convex parts.

5. Acquisition Setup

The RTI data has been acquired by a custom light dome with a radius of about 30*cm*, and with 48 LED lights placed as shown in Figure 2: 18 are put equally spaced in the azimuth coordinates at the elevation of 10 degrees, 12 at 30 degrees, 9 at 50 degrees, 8 at 70 degrees, and one parallel to the camera view direction. The LEDs are neutral white lights that covers the entire visible spectrum. The capture device is a 36.3 Mpixels DSLR FX Camera Nikon D810, and a 50*mm* AF Nikkor Lens. The acquisition system has been calibrated with a white planar target and reflective dark spheres, coupled with a light direction/intensity estimation procedure [CPM*16, GCD*18].
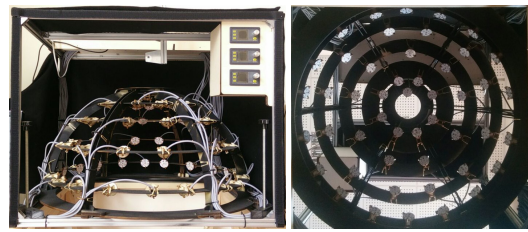

Figure 2: *Custom light dome used for the objects' acquisition.*

6. RTI Fitting

Although our evaluation framework can be applied to validate any RTI relighting method, here we focus on three techniques, i.e., two of the most used RTI data fitting approaches (PTM, HSH) and one of the most recent appearance profile interpolation strategies (RBFs). We do not cover multi-spectral processing, since the treatment of multi-channel, color signals is an orthogonal problem. After the acquisition of the raw RTI image stack, we thus pre-process the data to transform the color space into a *LRGB* representation, which is used in all the three fitting procedures. So, we consider a single chroma value per pixel and we apply the computation to the luminance channel. This choice also helps in reducing chromatic aberrations, mostly in metallic samples. We used in our comparison a second order PTM (6 coefficients), the de-facto standard representation used in many practical applications, and a third order HSH (16 coefficients) that is probably the most accurate global function commonly used for relightable images. We also use a 5-neighbors RBF as an alternative to those global fitting methods.

Polynomial Texture Maps. Polynomial Texture Maps (PTMs) [MGW01] are the current standard in RTI applications (mostly in the CH field) where a lot of museums use them to convey virtual, remote visualization of artworks. This is the seminal method in this field, and it proposed a per-pixel bi-quadratic polynomial interpolation of RTI stacks to produce

a spatially varying reflectance I as functions of light directions. At each pixel, given the six parameters α_i computed from data fitting, and a light direction $\vec{l}$, reflectance is computed as: $I\left(\vec{l}\right) = \alpha_0 \vec{l}_x^2 + \alpha_1 \vec{l}_y^2 + \alpha_2 \vec{l}_x \vec{l}_y + \alpha_3 \vec{l}_x + \alpha_4 \vec{l}_y + \alpha_5$. Due to the easiness in data capture, the simple model, which allows for a quite fast processing and efficient storing, and the availability of open source tools for RTI data processing and for web-publishing the resulting relightable images, PTM became the most widespread RTI visualization method. For this reason, we include it in our evaluation. For our test, we compute the six PTM parameters α_i by using the standard least square approach.

Hemispherical Harmonics. Similarly to the Fourier series, spherical harmonics are a good set of basis to describe functions on the surface of a sphere, and low order set of those basis are typically used in modeling low frequency reflectance, e.g., applied in photometric stereo scenarios [BJK07]. In the RTI acquisition setup we have a single viewpoint and only the upper hemisphere of surface normals is visible, rather than the whole sphere, and the spherical harmonics functions are no longer orthonormal. For this reason, the hemispherical basis has been introduced [MMC*08] to represent image irradiance, and they are defined from the shifted associated Legendre polynomials [ERF11]. They are capable to model near and asymptotically distant illumination conditions, and the resulting relighted images from new virtual light sources proved to better preserve contrast from the original images. For our test, we compute HSH parameters by using the standard Singular-Value Decomposition method. It has been proved that good results are obtained with sixteen parameters α_l^m per-pixel [Mac14], which consist in four first-order, five second-order and seven third-order terms. The per-pixel reflectance is computed as: $I\left(\vec{l}\right) = \sum_{l=0}^{n-1} \sum_{m=-l}^{l} \alpha_l^m H_l^m\left(\vec{l}\right)$, where H_l^m are the hemispherical harmonic basis functions.

Radial Basis Functions. Since the typical RTI acquisition produces a sparse set of samples within the (l_x, l_y) light direction space, Radial Basis Function (RBF) [Buh03] can be considered as an efficient way to interpolate them and produce a relighted image with a new virtual light [GCD*17]. Given N input images in the RTI capture, RBF interpolation is achieved by computing the parameters that define a sum of N radial functions. Here we use Gaussians centered at each light direction, with standard deviation R compute from the light distribution: the more are light directions, the smaller will be the value of R. In order to perform the fitting and find the parameters, for each radial function the light directions the closest five lights are considered to solve the fitting. Of course, this choice is due to a trade-off between local, high-frequency signal preservation (e.g., highlights and shadows), and proper data smoothing. Once all the per-pixel parameters α_i have been computed, the reflectance I is expressed as $I\left(\vec{l}\right) = \sum_{i=1}^{N} \alpha_i e^{\frac{\|\vec{l}-\vec{l}_i\|^2}{R^2}}$.

7. Objective Evaluation

The first part of the evaluation is devoted to an objective measurement of differences between the expected result of a light-based interpolation method and the actual outcome of the analyzed RTI techniques. The standard way to compute multi-light fitting quality is the so called leave-one-out strategy. Given an RTI raw stack

of N images, we loop on that set, and for each image we compute fitting parameters with $N-1$ images, by excluding the current image from the computation. Then we use those parameters to produce the relighted image corresponding to the excluded one. The error metric is computed by comparing the original and relighted images. We use two classical image comparison metrics, i.e., Peak Signal to Noise Ratio (PSNR) [WTFE18a] and Structural Similarity (SSIM) [WTFE18b], which have been recently used as well for measuring the compression quality in RTI data visualization [PCS18]. We compute those error metrics for all the three tested techniques (PTM, HSH, RBF) and the four objects of study (Coin1, Coin2, Lamina, Shell).

Method	PSNR				SSIM			
	Avg.	Med.	1st Qr	3rd Qr.	Avg.	Med.	1st Qr	3rd Qr.
PTM	22.20	23.10	19.43	25.77	0.69	0.71	0.59	0.79
HSH	23.82	24.17	**21.60**	26.87	0.73	0.75	0.65	0.82
RBF	**24.14**	**24.35**	21.00	**27.56**	**0.80**	**0.83**	**0.73**	**0.88**

Table 1: *Global PSNR and SSIM. Global statistics of PSNR and SSIM computed across all the four RTI image stacks of datasets* Coin1*,* Coin2*,* Lamina*, and* Shell*.*

To analyze the statistics of PSNR and SSIM, we present a series of error data measurements, i.e., the average, the median values, and the 25th and the 75th percentiles. In Table 1, we present the global error across all datasets. We can see that, on average, RBF is consistently better than the two other techniques, especially for the more perceptual metric SSIM. PTM is, by contrast, consistently presenting lower results. On the whole image set, the difference between RBF and HSH is not statistically significant at the 95% confidence level (even if close, $p = 0.07$ in a paired t-test for PSNR), while the differences between PTM and the other methods are both highly significant ($p < 10^{-14}$ in a paired t-tests for PSNR).

A better interpretation of the results can be given by looking at those statistics on a per-object basis (Table 2 and 3). The first three objects (i.e., *Coin1*, *Coin2*, and *Lamina*) exhibit a glossy material behavior. We can see how the RBF interpolation better conveys this metallic appearance. The per-image PNSR does not improve so much because the highlight region is typically small compared to the entire image, and due to the fact that the neighborhood based RBF-sampling in the leave-one-out framework results in a slight drift of the rendered highlight peaks. Conversely, under a more perceptual metrics (i.e., SSIM), we can observe a significant increase in visual performances while relighting those three metallic objects. In particular, while a big number of images are almost diffuse for the coins, in the *Lamina* data a large part of the per-pixel appearance profile has highlights, and we can see here how the improvement is bigger by using the local-driven fitting of the RBF approach. An interesting result arises from the matte object *Shell*. Here, the HSH method proves to be a better choice in rendering new virtual light source positions. The reason is that with a diffuse, low-frequency optical behavior the choice of a global fitting method (as HSH), that takes into account the entire appearance profile, is capable to reasonably model the regularity of behavior. RBF, which weights only the contribution of a small number of close lights, is, instead, introducing more error in the areas near shadows. Looking at the statistical significance of the differences, it is possible to understand the variability of the methods' behavior on different materials. For the shell, there is a significant difference between

Method	PSNR															
	Coin1				Coin2				Lamina				Shell			
	Avg.	Med.	1st Qr	3rd Qr.	Avg.	Med.	1st Qr	3rd Qr.	Avg.	Med.	1st Qr	3rd Qr.	Avg.	Med.	1st Qr	3rd Qr.
PTM	20.96	22.37	17.14	25.56	22.43	23.71	19.85	25.60	20.27	19.67	16.03	24.26	24.53	24.78	21.77	27.50
HSH	22.92	23.01	21.37	26.1	23.67	24.00	22.07	26.11	21.26	20.99	16.63	25.44	**26.60**	**26.68**	**24.41**	29.51
RBF	**23.74**	**24.83**	**21.99**	**27.22**	**24.35**	**25.01**	**22.37**	**26.96**	**22.45**	**21.01**	**16.88**	**27.90**	25.48	24.23	21.61	**29.82**

Table 2: *PSNR vs Objects PSNR statistics corresponding to PTM, HSH, and RBF relighting techniques for each of the studied objects, i.e.,* Coin1, Coin2, Lamina, *and* Shell.

Method	SSIM															
	Coin1				Coin2				Lamina				Shell			
	Avg.	Med.	1st Qr	3rd Qr.	Avg.	Med.	1st Qr	3rd Qr.	Avg.	Med.	1st Qr	3rd Qr.	Avg.	Med.	1st Qr	3rd Qr.
PTM	0.61	0.64	0.49	0.68	0.70	0.73	0.65	0.75	0.61	0.59	0.50	0.71	0.81	0.83	0.75	0.89
HSH	0.66	0.68	0.61	0.75	0.75	0.76	0.71	0.80	0.61	0.64	0.53	0.72	**0.85**	**0.87**	**0.82**	0.91
RBF	**0.77**	**0.82**	**0.70**	**0.87**	**0.81**	**0.84**	**0.76**	**0.88**	**0.78**	**0.80**	**0.72**	**0.88**	0.81	0.83	0.72	**0.93**

Table 3: *SSIM vs Objects SSIM statistics corresponding to PTM, HSH, and RBF relighting techniques for each of the studied objects, i.e.,* Coin1, Coin2, Lamina, *and* Shell.

HSH and PTM ($p < 10^{-5}$ in a paired t-test for PSNR) and an almost significant difference between HSH and RBF ($p = 0.02$ in a paired t-test for PSNR) and between RBF and PTM ($p = 0.03$ in a paired t-test for PSNR). On the other hand, for the lamina, we get a significant improvement using RBF rather than HSH ($p = 0.03$ in paired t-test for PSNR) or PTM ($p < 10^{-5}$ in paired t-test for PSNR). The difference between PTM and HSH is not statistically significant ($p = 0.07$). For the coins, RBF based relighting are always significantly better than HSH based ($p = 0.005$ for Coin 1 and $p = 0.004$ for Coin 2 in paired t-tests) and HSH based relighting are always significantly better than PTM based ones ($p < 10^{-7}$ for Coin 1 and $p < 10^{-5}$ for Coin 2 in paired t-tests for PSNR). Similar trends are obtained testing SSIM differences.

In Figure 3 and 4 we respectively plot the PSNR and SSIM as functions of light angle of incidence. Since we would like to understand how the fitting behaves in the presence of specular reflection, we put in the x-axis the *Half Angle*, i.e., the angle between the half vector and the surface normal. Given L the direction of the light, and V the view vector, the half vector will be $H = (L+V) / \|L+V\|$; since we have almost flat objects we consider the average per-image normal as $\vec{n} = [0, 0, 1]$. For metallic objects we can see how the numerical error in the relighting strongly depends on this angle, which, for a dome measuring an almost flat object, is roughly proportional to light elevation. As soon as we get to the specular angular region the quality of the rendering decreases (Figure 3a, 3b, and 3c). For a diffuse object (Figure 3d the PSNR increases; in fact, in this case, we have more non-idealities for raking lights than for lights aligned with the mirror-reflection direction (e.g., self and cast shadows). On the other hand, while the perceptual SSIM is statistically better for RBF than the other techniques, its behavior seems to be independent of the half angle (Figure 4).

8. Subjective Evaluation

In order to evaluate the effectiveness of the three chosen RTI visualization techniques in terms of performance, suitability for the CH field, and the improvement of the visual quality of rendered objects, we performed a thorough user evaluation involving quantitative and subjective measurements, by conducting an online survey through a web-based questionnaire carried out by a number of volunteers.

Goal. The main goal of the evaluation was to assess which RTI fitting and rendering method is more adequate for the usage in the typical scenario of object inspection and daily research activity, where many users with different skills and experiences try to interactively explore virtually relighted artworks. Given the large number of fitting and relighting methods, it is out of the scope of this work to provide a fully comprehensive comparison of existing techniques and outcomes. As explained in previous sections, similarly, in this user-based evaluation we limited our comparison to the three RTI fitting techniques presented before, i.e., PTM, HSH, and RBF.

Setup. The experimental setup considered the RTI data of the objects presented in section 4, and the fitting methods explained in section 6. For each RTI stack we consider a subset of five representative images taken at different light source elevations, each corresponding to a different parallel of lights in the dome; in this way we can test the visual performances in the presence of shadows (i.e., raking lights) or highlights. For each of those original images we compute the relighted one in the same leave-one-out sense (Sec. 7) for the PTM, HSH, and RBF techniques. At the end of this procedure we will have five quadruples (original, PTM, HSH, and RBF image) for each dataset; in total, we will use a set of 76 images for our visual evaluation. In Figure 5 we show one quadruple from each analyzed object.

Tasks. The experiments consist in two main sections/tasks, which ask respondents to give preferences on the relighted images based on their experience in daily artwork inspection, and the visual quality of CH item presentation. The first test asks respondents to select the most similar relighted images with respect to ground truth data. We show the user a sequence of three images (Figure 6 Top). The centered one is a photo of the real artwork (ground truth image). The lateral ones are two computer-generated images of the same artwork rendered from the same point of view of the ground truth image and relighted with the same light source intensity and direction of the ground truth image. The user has to click/select the lateral image whose rendering he considers the most similar compared to the center ground-truth image. The second one is not guided by the presence of a real photograph, and asks respondents to provide an "unsupervised" preference between two RTI virtual relightings (Figure 6 Bottom). We show the volunteer a sequence of two computer-generated images of the same artwork viewed by the same point and relighted with the same light source intensity and direction. He has to select the one whose rendering he prefers as a meaningful visualization related to his daily CH inspection activity and quality expectation.

Participants. 25 participants were recruited across various leading institutions involved in CH studies and applications (see Ac-
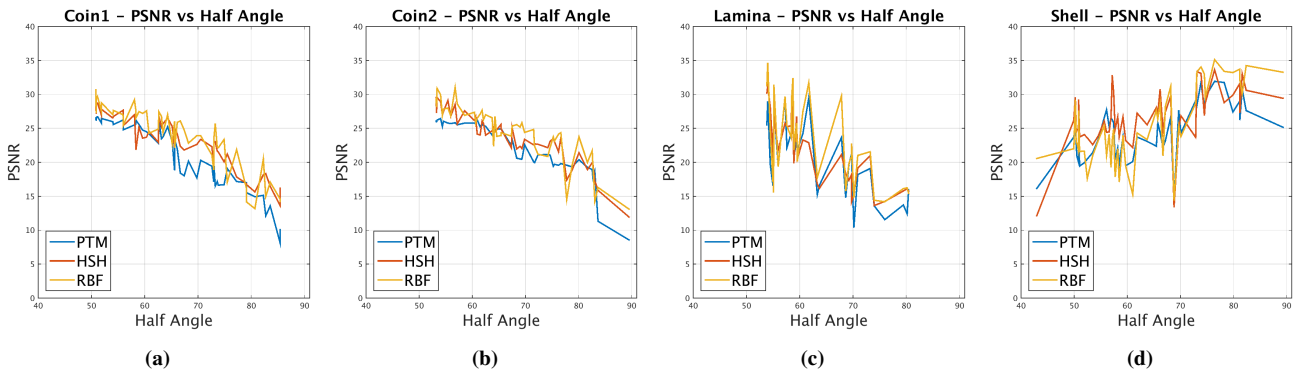
(a)　　　　　　　　(b)　　　　　　　　(c)　　　　　　　　(d)

Figure 3: *PSNR vs Half Angle PSNR statistics corresponding to PTM, HSH, and RBF relighting techniques as functions of the half angle (the angle between the half vector and the average per-image normal): (a) Coin1; (b) Coin2; (c) Lamina; (d) Shell.*
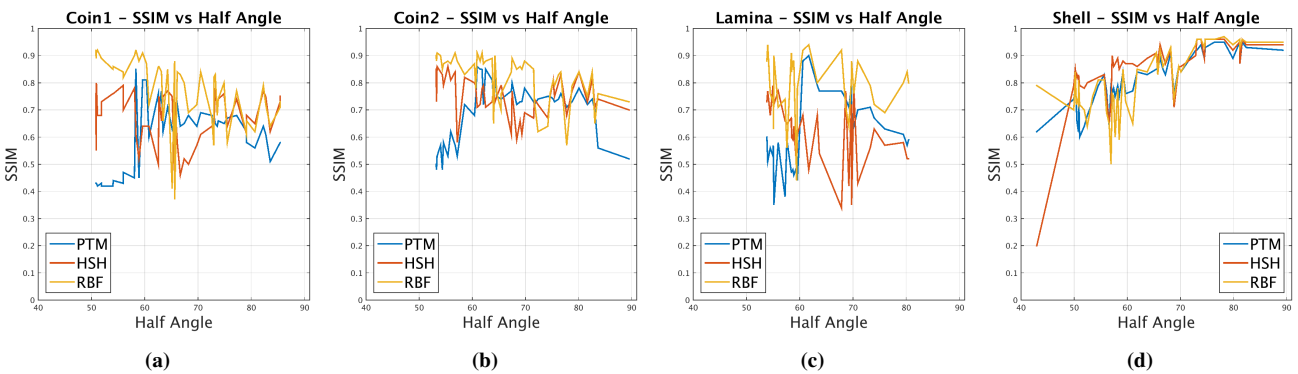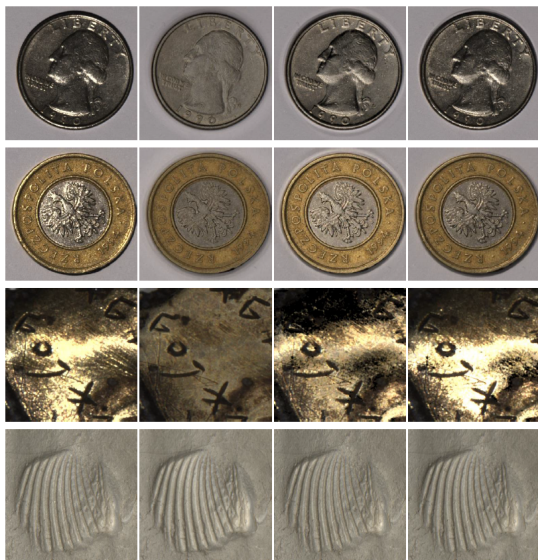

(a)　　　　　　　　(b)　　　　　　　　(c)　　　　　　　　(d)

Figure 4: *SSIM vs Half Angle PSNR statistics corresponding to PTM, HSH, and RBF relighting techniques as functions of the half angle (the angle between the half vector and the average per-image normal): (a) Coin1; (b) Coin2; (c) Lamina; (d) Shell.*

Figure 5: *Objects. Rows (top to bottom): coin1, coin2, lamina and shell. Columns (left to right): original image, PTM, HSH, and RBF relightings.*

Figure 6: *Web-based Questionnaire. Two tests have been performed in a web-based questionnaire. The first (top figure) is a visual comparison guided by using a ground truth image, while the second (bottom figure) asks an "unsupervised" preference between two virtual relighted images.*

knowledgments). They are subdivided in two groups, the experts working in different areas related to the CH field, such as art historian/curators, restorer/conservators, conservation scientists (12 in total), and non-experts volunteers (another 13 people). Considering
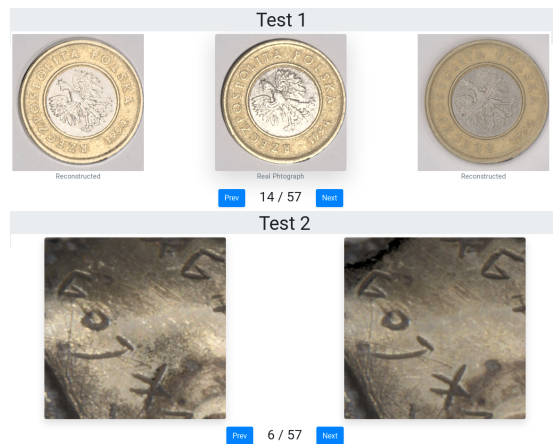
those two groups is due to the fact that RTI rendering and relighting has been widely used not only for CH studies performed by expert scholars, but also for cultural dissemination and virtual presentation to the more general public. All the tests are as much blind as possible; all the respondents are unaware of the type of relighting methods we have used to construct the relighted images, and they
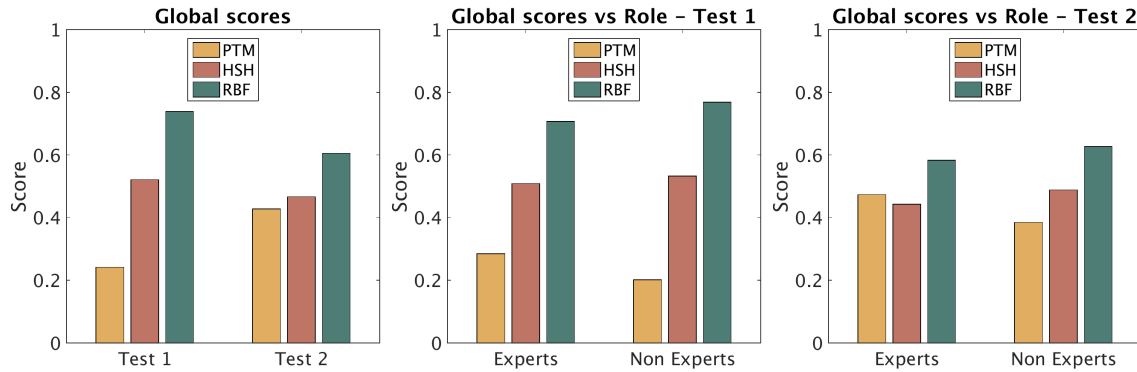
Figure 7: *Global Scores. Global number of votes for PTM, HSH, and RBF techniques across all datasets. We present statistics for* Test 1 *and* Test 2 *(Left). We also present the same global votes splitted by type of participants, i.e., CH people and others (Middle and Right).*
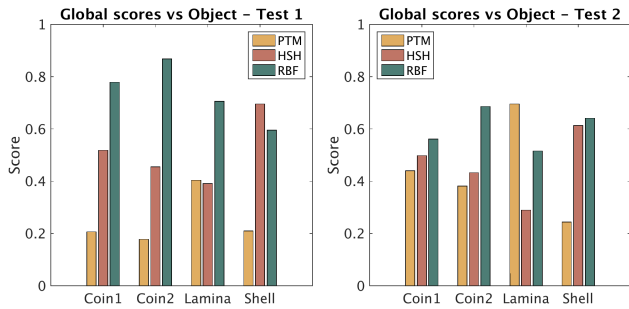

Figure 8: *Scores vs Objects. We present the scores of the three techniques for each studied object. RBF wins for metallic samples, while HSH and RBF are comparable for matte materials. As in Figure 7, the missing ground truth reference in* Test 2 *flatten the overall scores.*
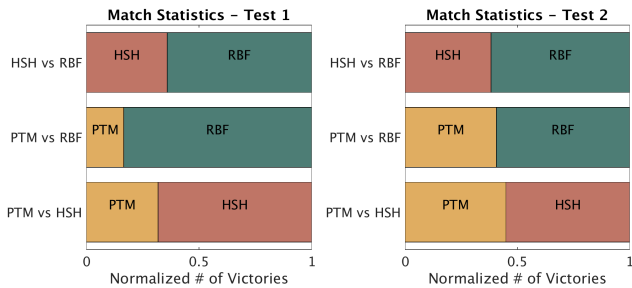

Figure 10: *Lamina. An image for the subjective evaluation. Top row: original image and PTM relighting. Bottom row: HSH and RBF relightings.*

Figure 9: *Match Statistics. Matches between the techniques (PTM vs HSH, PTM vs RBF, and HSH vs RBF) are displayed for* Test 1 *(Left) and* Test 2 *(Right). Values are normalized number of victories across all objects.*

have never worked on those objects before. This helps us to obtain an unbiased evaluation of the RTI fitting techniques and their results on the objects of study.

Design. We carried out the evaluation using a website developed with *HTML5* and *Javascript*. Users access the test by a private link. Before starting with the two tests described before, participants are asked to fill a small, anonymous form with four questions about their institution, role, usual tasks and kind of objects they use to work with. Then a few lines explain every test, where users can click the left or right image, or none of them. Even if the whole combination of pair images is exposed to all the users, the showing sequence was randomized for every session, as well as the left-right order of the comparisons. It is also allowed to go forwards

and backwards through the pairs of images, and change the elected one in every moment. We set limitless time for the experiment, as we want users to inspect and select their options carefully and have the chance to correct previous elections. Instead, we set a limit of not-null answers of 50% for every test. The average filling time stands around 14 minutes.

Performance evaluation. The evaluation of the subjective quality of the relightings is based on a "score" given by each subject to each technique on each original image and for each of the two tests. This score is simply the number of "wins" in the binary comparisons, that can be, clearly, 0, 1, or 2. We divide this score by two to obtain a normalized value and perform statistics over the different tests, user profiles and objects. In a few cases, where the user did not perform all the three binary comparisons for the same relighting, we discarded the score from the statistics.

In the objective evaluation we have witnessed how PTM is the technique that behaves poorly across datasets compared with the other two; HSH improves a little, by trying to render better non-diffuse material properties, while RBF on average wins over both other techniques. Now, we are interested in understanding if this
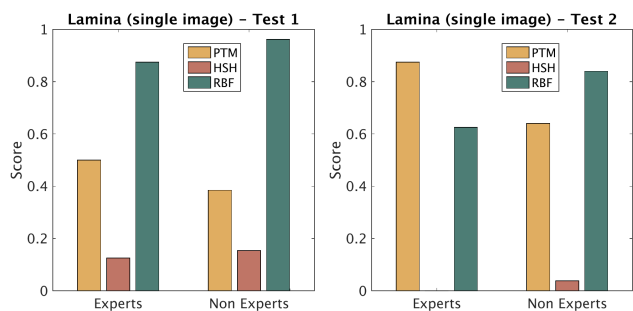
Figure 11: *Lamina - Single Image Scores. Voting scores for the* Lamina *image in Figure 10. While scores in* Test 1 *(ground truth comparison) are similar between experts and non-experts,* Test 2 *produces non consistent results between the two groups. This is probably due to the fact that experts, without a ground truth, prefer a more readable, smoothed image (i.e., PTM), while non-experts tend to vote to a more photorealistic one (i.e. RBF).*

result has been confirmed visually and perceptually by the users. Figure 7 shows the scores of the three techniques for the first and the second tests, averaged over the voting subjects, objects and light directions. We can see how RBF wins in both tests. An interesting result is that it wins more when the user judgment is guided by the presence of the ground truth image, while PTM scores improve in the "unsupervised" second test. This is probably due to the fact that PTM removes many sharp hilights, a non-photorealistic approach which might be preferred in some cases, even though far from ground truth. Checking the statistical significance of the differences, we see that, for *Test 1* pairwise Wilcoxon signed rank tests between couple of methods provide always highly significant differences ($p < 10^{-20}$ comparing PTM and HSH, $p < 10^{-56}$ comparing PTM and RBF, $p < 10^{-15}$ comparing HSH and RBF). There is a clear ranking between the perceived quality obtained with the different methods. Statistical tests also confirm the fact that the quality of PTM is no longer considered so bad in *Test 2*. There is still a clearly significant difference between groups, but only pairwise differences between RBF and HSH ($p < 10^{-4}$ in Wilcoxon test) and between PTM and RBF ($p < 10^{-6}$ in Wilcoxon test) are significant, while the difference between PTM and HSH is not ($p = 0.65$ in Wilcoxon test). By separately analyzing votes of experts and non-experts other interesting differences emerge. The results of *Test 1*, "guided" by the displayed ground truth images, are consistent between those two groups, while in the second test CH people don't make a clear choice on the best method. Even with the presence of burned highlights, and also of some artifacts, non-experts consider that the renderings with more highlights and shadows are better, since they probably judge them more photorealistic, even without a looking at a ground truth image (the behavior is consistent with the *Test 1*). Conversely, in the *Test 2*, the CH people tend to increase the preference on a smoothed PTM signal, which, although it fails to render high-frequency material behavior, it conveys a cleaner (and readable) surface visualization. An example of this statistical behavior is given in Figure 10 and Figure 11. There we show only one image of the dataset *Lamina* and we present the original photo, and the PTM, HSH, and RBF rendering. From the statistics corresponding to just this photo and its relightings, we can see how in the *Test 1* RBF wins both for experts and non-expert people, while in the *Test 2* the votes are not consistent between the

two groups, i.e., PTM wins for experts while RBF has more votes for non-experts. Statistical tests show that, while all the pairwise differences between groups are highly significant for Test 1 (for experts we get $p < 10^{-8}$ comparing PTM and HSH, $p < 10^{-25}$ comparing PTM and RBF, $p < 10^{-5}$ comparing HSH and RBF in Wilcoxon signed rank tests, for non experts $p < 10^{-13}$ comparing PTM and HSH, $p < 10^{-32}$ comparing PTM and RBF, $p < 10^{-9}$ comparing HSH and RBF), but this is no longer true for Test 2. For both experts and non-experts, the difference between PTM and HSH is not statistically significant ($p = 0.23$ for experts, $p = 0.07$ for non-expert). For the experts the differences between HSH and RBF and PTM and RBF are significant only at the 95% confidence level ($p = 0.013$ and $p = 0.032$ in Wilcoxon tests). For non experts we have, instead, highly significant differences ($p < 10^{-2}$ and $p = < 10^{-6}$ in Wilcoxon tests). Figure 8 presents the same results for each single object. They confirm two trends: in *Test 1* the RBF wins over the other techniques, while in *Test 2* the differences between preferences are flattened. In the matte object *Shell* HSH and RBF are equally voted; this is not surprising, since a local fitting method such as RBF is expected to behave better with high-frequency signals (e.g., highlights). The different trends are confirmed by statistical analysis: considering *Test 1*, we have that, for the matte object (shell), the pairwise comparison does not show significant differences between HSH and RBF ($p - value = 0.06$ in Wilcoxon test) while HSH/PTM and RBF/PTM differences are highly significant ($p < 10^{-4}$ and $p < 10^{-7}$, respectively). For the highly specular object (lamina), instead, there is no statistically significant difference between PTM and HSH ($p - value = 0.93$ in Wilcoxon test), while the differences between RBF and the other methods are highly significant ($p < 10^{-6}$ comparing PTM and RBF, $p < 10^{-4}$ comparing HSH and RBF). For the intermediate behavior (coins), the differences between all the groups are highly significant. RBF is always clearly the better technique. In *Test 2*, for the shell we get the same results of *Test 1*: p-values are quite low in comparisons between HSH/PTM ($< 10^{-10}$) and RBF/PTM ($< 10^{-11}$), but there is no statistically significant difference between HSH and RBF ($p = 0.80$ in Wilcoxon test). The outcomes of Test 2 are, however, different for the Lamina, where the ranking of the methods is different with respect to Test 1 and differences are always statistically significant: PTM wins more than the other methods and the p-values are statistically significant ($p < 10^{-8}$ in comparison with HSH and $p < 10^{-5}$ in comparison with RBF), and there is also a statistically significant preference of RBF over HSH ($p < 10^{-4}$). The reason is probably due to the failure of the HSH relighting for high level of specularity and the better visibility of details in non-specular relighting obtained with PTM. For the coins, we have that the methods seems equivalent on Coin 1, where there are no statistically significant differences between the methods, this may be due to the fact that without references some users prefer "matte" rendering and others a specular one. On Coin 2, however, we have statistically significant preferences for RBF relighting both with respect to PTM ($p < 10^{-5}$ in Wilcoxon test) and HSH ($p < 10^{-4}$). In Figure 9 Left, we show how, in *Test 1*, the HSH wins over the PTM, the RBF wins more against the PTM, and how RBF exhibits a better quality compared to HSH. Similarly, single matches in the *Test 2* (Figure 9 Right) confirm the trend that in the "unsupervised" test the scores are flattened; nevertheless, the RBF proved to be again the best technique.

9. Discussion

Analyzing the quality of relightings from RTI stacks is interesting as their analysis is commonly used for objects' inspection in CH. So far, issues related to the quality of the renderings have not been addressed in the literature. Our comparisons of different methods on different materials give useful insights. RTI is extremely popular in CH object inspection due to its ease of capture and capability to quickly produce relightable images. However, the relatively sparse sampling of the light direction field and the simplicity of the fitting/interpolation approaches may lead to errors with complex objects and appearance properties. On matte surfaces, as the *Shell*, synthetic images are closer to the reference ones, and the technique used is less critical. Greater errors appear with raking lights as the projected shadows are not reproduced accurately. On challenging specular surfaces (coins, lamina), the quality is lower as expected. PSNR clearly decreases with elevation, while SSIM is not similarly changed. This may be caused by large-scale luminance shifts between original and relighted images, compensated by the perceptive adjustments of the SSIM metric. While the RBF interpolation seems to provide the best similarity with respect to original images both in objective and subjective tests (and 3rd order HSH is clearly superior to second order PTM), CH experts been asked to compare relighted images without reference do not show clear preferences (Figure 7 and 10), showing that non-photorealistic depictions might be accepted if judged more readable. These results lead to the identification of two competing research directions. First of all, it looks interesting to devise specialized task-specific relighting techniques, which do not strive to photorealistically reproduce the original object, but limit themselves to reproduce the set of interesting features for a particular task (e.g., matte behavior for surface readability, discontinuities for crack detection). Since fitted representations are shown to be significantly lossy, these depictions should be extracted from raw data, rather than as enhancements of fitted models. On the other hand, to provide a photorealistic rendering, while still requiring only a relatively sparse set of input images, it might prove interesting to introduce some level of domain-specific knowledge in the light interpolation process, while still avoiding to solve the complex problem of full geometry and appearance reconstruction. Due to the presence of many BRDF datasets and benchmarks, and the possibility of producing a huge quantity of controlled relightings of different materials and geometries, an emerging research path is to exploit learning approaches to extract information from very sparse input samples [DAD*18].

Our preliminary analysis has clearly also some limitations. First, it is based on static images. Interactive variation of the light direction allows a better understanding of surface and material details, especially if the interpolation methods can capture well highlights and shadows. So, CH experts could probably appreciate even more HSH or RBF interpolation in the case of interactive relighting of non-matte objects. Testing the quality of interactive relighting is, however much more difficult, due to the necessity of designing specific tasks. We plan to perform this investigation in future work. Second, it is based on high-quality relighting made with floating point coefficients and without compression. Interactive relighting is often performed, in practice, by using data structures storing approximate values (often converted in 8 bit depth) and using compressed images. On-line relighting with RBF has been proposed using PCA compression of the image stack further reducing accuracy. Of course, it could be interesting to evaluate the effects of compression on the results. Another issue to be considered is related to the metrics used for quantitative comparison, that are generic and averaged on the whole image. It may be interesting to develop specifically designed to capture the ability to preserve a specific feature of interest for the CH analysis.

10. Conclusions

We have presented an evaluation framework to test three RTI-based rendering and relighting techniques, i.e., PTM, HSH, RBF. We used the data acquired with a fixed light dome, and we present both an objective and subjective evaluation scenario. The first tests the quality of reconstructed images through different error metrics (i.e., PSNR and SSIM) and by using a leave-one-out strategy. The subjective test has involved both experts and non-experts volunteers for a visual test of the quality of virtually relighted images. Although our work is a preliminary attempt to define a standard way to judge the perceived quality of RTI-based visualization, we already found that the objective and user-guided outcomes behave consistently. A lot of interesting work should be done to further investigate these results. So far, both tests proved that local interpolation methods (RBF) are more convenient than more standard techniques (PTM/HSH) to render high-frequency details.

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

[AIK13] ARTAL-ISBRAND P., KLAUSMEYER P.: Evaluation of the relief line and the contour line on greek red-figure vases using reflectance transformation imaging and three-dimensional laser scanning confocal microscopy. *Studies in Conservation 58*, 4 (2013), 338–359. 1

[Bar65] BARRECA F.: Nuove iscrizioni fenicie da Sulcis. *Oriens Antiquus 4*, 1 (1965), 1–5. 3

[BCDG13] BROGNARA C., CORSINI M., DELLEPIANE M., GIA-CHETTI A.: Edge detection on polynomial texture maps. In *Proc. ICIAP* (2013), pp. 482–491. 2

[BJK07] BASRI R., JACOBS D., KEMELMACHER I.: Photometric stereo with general, unknown lighting. *International Journal of Computer Vision 72*, 3 (2007), 239–257. 4

[BSMG05] BARTESAGHI A., SAPIRO G., MALZBENDER T., GELB D.: Three-dimensional shape rendering from multiple images. *Graphical Models 67*, 4 (2005), 332–346. 2

[Buh03] BUHMANN M. D.: *Radial basis functions: theory and implementations*, vol. 12. Cambridge university press, 2003. 4

[CHI17] CHI: Cultural heritage imaging, 2017. [Online; accessed 25-June-2018]. URL: http://culturalheritageimaging.org. 2

[CPM*16] CIORTAN I. M., PINTUS R., MARCHIORO G., DAFFARA C., GIACHETTI A., GOBBETTI E.: A practical reflectance transformation imaging pipeline for surface characterization in cultural heritage. In *Proc. GCH* (2016). 3

[DAD*18] DESCHAINTRE V., AITTALA M., DURAND F., DRETTAKIS

G., BOUSSEAU A.: Single-image svbrdf capture with a rendering-aware deep network. *ACM TOG 37*, 128 (2018), 15. 9

[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*, 4-5 (2012), 317–331. 2

[Dix13] DIXON H. M.: *Phoenician Mortuary Practice in the Iron Age I-III (ca. 1200-ca. 300 BCE).* PhD thesis, U. Michigan, 2013. 3

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

[Duf10] DUFFY S.: Polynomial texture mapping at Roughting Linn rock art site. In *Proc. ISPRS Commission V Mid-Term Symposium: Close Range Image Measurement Techniques* (2010), pp. 213–217. 1

[Duf13] DUFFY S. M.: *Multi-light imaging for heritage applications.* English Heritage, 2013. 2

[ERF11] ELHABIAN S. Y., RARA H., FARAG A. A.: Towards accurate and efficient representation of image irradiance of convex-Lambertian objects under unknown near lighting. In *Proc. ICCV* (2011), pp. 1732–1737. 4

[FAR07] FATTAL R., AGRAWALA M., RUSINKIEWICZ S.: Multiscale shape and detail enhancement from multi-light image collections. *ACM TOG 26*, 3 (2007), 51. 2

[GCD*17] GIACHETTI A., CIORTAN I., DAFFARA C., PINTUS R., GOBBETTI E.: Multispectral RTI analysis of heterogeneous artworks. In *Proc. GCH* (2017). 2, 3, 4

[GCD*18] GIACHETTI A., CIORTAN I., DAFFARA C., MARCHIORO G., PINTUS R., GOBBETTI E.: A novel framework for highlight reflectance transformation imaging. *Computer Vision and Image Understanding 168* (2018), 118–131. 3

[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 *Optics for Arts, Architecture, and Archaeology V* (2015), vol. 9527, p. 95270B. 2

[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. 2

[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

[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. 2

[KK13] KOTOULA E., KYRANOUDI M.: Study of ancient greek and roman coins using reflectance transformation imaging. *E-conservation magazine 25* (2013), 74–88. 1

[Leu15] LEUVEN K.: Multispectral microdome, 2015. [Online; accessed 25-June-2018]. 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. *IEEE TPAMI 36*, 1 (2014), 86–98. 2

[LLW12] LAM P.-M., LEUNG C.-S., WONG T.-T.: Noise-resistant hemispherical basis for image-based relighting. *IET image processing 6*, 1 (2012), 72–86. 2

[Mac14] MACDONALD L. W.: Colour and directionality in surface reflectance. In *Proc. AISB* (2014). 4

[Mac15] MACDONALD L. W.: *Realistic visualisation of cultural heritage objects.* PhD thesis, UCL, 2015. 2

[MBW*14] MANFREDI M., BEARMAN G., WILLIAMSON G., KRONKRIGHT D., DOEHNE E., JACOBS M., MARENGO E.: A new quantitative method for the non-invasive documentation of morphological damage in paintings using RTI surface normals. *Sensors 14*, 7 (2014), 12271–12284. 1

[MGW01] MALZBENDER T., GELB D., WOLTERS H.: Polynomial texture maps. In *Proc. ACM SIGGRAPH* (2001), pp. 519–528. 2, 3

[MMC*08] MUDGE M., MALZBENDER T., CHALMERS A., SCOPIGNO R., DAVIS J., WANG O., GUNAWARDANE P., ASHLEY M., DOERR M., PROENCA A.: Image-based empirical information acquisition, scientific reliability, and long-term digital preservation for the natural sciences and cultural heritage. In *Eurographics (Tutorials)* (2008). 2, 4

[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. 6, pp. 29–40. 1

[Pan16] PAN R.: Detection of edges from polynomial texture maps. *3D Research 7*, 1 (2016), 3. 2

[Pay13] PAYNE E. M.: Imaging techniques in conservation. *Journal of conservation and museum studies 10*, 2 (2013). 2

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

[PCS18] PONCHIO F., CORSINI M., SCOPIGNO R.: A compact representation of relightable images for theweb. In *Proc. ACM Web3D* (2018), ACM Press, p. 10. 2, 4

[PGPG17] PINTUS R., GIACHETTI A., PINTORE G., GOBBETTI E.: Guided robust matte-model fitting for accelerating multi-light reflectance processing techniques. In *Proc. BMVC* (2017). 2

[PLGF*15] PITARD G., LE GOÏC G., FAVRELIÈRE H., SAMPER S., DESAGE S.-F., PILLET M.: Discrete modal decomposition for surface appearance modelling and rendering. In *Optical Measurement Systems for Industrial Inspection IX* (2015), vol. 9525, p. 952523. 2

[PPY*16] PINTUS R., PAL K., YANG Y., WEYRICH T., GOBBETTI E., RUSHMEIER H.: A survey of geometric analysis in cultural heritage. In *Computer Graphics Forum* (2016), vol. 35, pp. 4–31. 2

[RL05] ROUSSEEUW P. J., LEROY A. M.: *Robust regression and outlier detection*, vol. 589. John wiley & sons, 2005. 2

[RTF*04] RASKAR R., TAN K.-H., FERIS R., YU J., TURK M.: Nonphotorealistic camera: depth edge detection and stylized rendering using multi-flash imaging. In *ACM TOG* (2004), vol. 23, pp. 679–688. 2

[SSWK13] SCHWARTZ C., SARLETTE R., WEINMANN M., KLEIN R.: Dome ii: A parallelized btf acquisition system. In *Material Appearance Modeling* (2013), pp. 25–31. 2

[STMI14] SHI B., TAN P., MATSUSHITA Y., IKEUCHI K.: Bipolynomial modeling of low-frequency reflectances. *IEEE TPAMI 36*, 6 (2014), 1078–1091. 2

[SWM*16] SHI B., WU Z., MO Z., DUAN D., YEUNG S.-K., TAN P.: A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo. In *Proc. ICCV* (2016), pp. 3707–3716. 2

[Sze10] SZELISKI R.: *Computer vision: algorithms and applications.* Springer Science & Business Media, 2010. 2

[TGVG12] TINGDAHL D., GODAU C., VAN GOOL L.: Base materials for photometric stereo. In *Proc. ECCV* (2012), pp. 350–359. 2

[UW13] URIBE M. D.-G., WHEATLEY D. W.: Rock art an digital technologies: the application of reflectance transformation imaging (RTI) and 3D laser scanning to the study of late bronze age iberian stelae. *Menga: Revista de prehistoria de Andalucía*, 4 (2013), 187–203. 1

[WGSD09] WANG O., GUNAWARDANE P., SCHER S., DAVIS J.: Material classification using brdf slices. In *Proc. CVPR* (2009), pp. 2805–2811. 2

[Woo80] WOODHAM R. J.: Photometric method for determining surface orientation from multiple images. *Optical engineering 19*, 1 (1980), 191139. 2

[WTFE18a] WIKIPEDIA, THE FREE ENCYCLOPEDIA: Peak signal-to-noise ratio, 2018. [Online; accessed 25-June-2018]. URL: https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio. 4

[WTFE18b] WIKIPEDIA, THE FREE ENCYCLOPEDIA: Structural similarity, 2018. [Online; accessed 25-June-2018]. URL: https://en.wikipedia.org/wiki/Structural_similarity. 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), 25. 2