A Modular Approach for Leveraging BRDF Monotonicity in Computing Shading Normals from Multi-Light Image Collections

RUGGERO PINTUS*, ANTONIO ZORCOLO*, and ENRICO GOBBETTI*, CRS4, Italy

Multi-Light Image Collections (MLICs) are stacks of photos acquired with a fixed viewpoint and varying surface illumination. They are typically used to provide rich information to characterize imaged objects. In particular, many simple and sophisticated Photometric Stereo solutions exist to produce detailed per-pixel normal maps that geometrically characterize the imaged model up to the finest details. These maps are commonly employed for per-pixel BRDF fitting to generate easy-to-use general shape and material representations for visual exploration. However, discrepancies between the chosen BRDF space , the complete optical behavior of complex objects, and the presence of non-local lighting effects in measurements may lead to suboptimal visual outcomes even with the most accurate geometric normal recovery. This article introduces a modular component designed to convert the geometric normals into well-behaved shading normals, under the common and general assumption that the reflectance under local illumination must be a monotonic function of the angle between the shading normal and the bisector of lighting and viewing directions. Since it does not require the coupling of shape and material estimation, the module allows seamless integration into existing reconstruction pipelines, supporting the mixing and matching of Photometric Stereo methods, BRDF models, and BRDF fitters. We demonstrate our approach's efficacy as a versatile solution for enhancing the fidelity of rendered images derived from MLICs through synthetic data, laboratory experiments, and real-world painting measurements.

CCS Concepts: • **Computing methodologies** → **Appearance and texture representations**; **Reflectance modeling**; *Scene understanding*.

Additional Key Words and Phrases: MLIC, Photometric Stereo, Reflectance computation, BRDF fitting, Monotonicity, Virtual Relighting, Paintings

ACM Reference Format:
Ruggero Pintus, Antonio Zorcolo, and Enrico Gobbetti. 2025. A Modular Approach for Leveraging BRDF Monotonicity in Computing Shading Normals from Multi-Light Image Collections. *ACM J. Comput. Cult. Herit.* 18, 0, Article Just Accepted (2025), 33 pages. https://doi.org/10.1145/3737292

1 Introduction

In recent years, the emergence of Multi-Light Image Collections (MLICs) has significantly expanded the horizons of computer vision and computer graphics. Composed of digital images taken from a fixed viewpoint but under diverse lighting conditions (e.g., positions, directions, wavelengths), they provide a rich source of information essential for comprehending and modeling the visual complexities of scenes and objects. A typical configuration involves capturing the appearance of each scene point under fixed view and various incident light directions only. This results in a comprehensive collection of samples (i.e., per-pixel appearance profiles) often organized into

*The authors contributed equally to this research.

Authors' Contact Information: Ruggero Pintus, ruggero.pintus@crs4.it; Antonio Zorcolo, antonio.zorcolo@crs4.it; Enrico Gobbetti, enrico.gobbetti@crs4.it, CRS4, Cagliari, Italy.

ACM 1556-4711/2025/0-ARTJust Accepted
https://doi.org/10.1145/3737292

image stacks. Due to these attributes, MLICs have found diverse applications across various domains, including Cultural Heritage (CH), natural science, industry, underwater investigation, medical imaging, and beyond [39].

The journey from the raw MLIC data to a complete, measurable, repeatable, reliable, and visually realistic representation of the object surface passes through a series of (often modular and standalone) computational steps that aim at explicitly characterizing the pure shape geometry and the optical response of each surface point. Besides a preliminary geometric and radiometric calibration of the MLIC data, the classical first step consists of modeling the geometry of the surface in terms of pixelwise normals (e.g., by using Photometric Stereo (PS) techniques [1]). This field is very active, and, through various advances in numerical solvers and emerging learning-based techniques (see Sec. 2.2), the quality of recovered normals is improving even for objects exhibiting very intricate geometry and complex material behaviors, offering a solid basis for many shape analysis applications [42]. To support visual exploration, these normals are coupled with a model of the optical behavior of the surface material, typically in terms of Spatially-Varying Bidirectional Reflectance Distribution Functions (SV-BRDFs). This is done, in most cases, by computing, per pixel, the coefficients of an analytical BRDF [18]. Having a normal+BRDF representation is very appealing since it can be easily distributed using low-bitrate representations, produces a physically reasonable result, and allows for natural integration with standard high-quality and real-time rendering solutions.

However, such a visual model may sometimes fail to match the original measurements well. This issue arises not only when geometric normals are inaccurately recovered by the Photometric Stereo method but also in cases of a "perfect" geometric reconstruction that is fully suitable for geometric inspection purposes. Several practical factors contribute to this problem. In addition to the consideration that input data may include non-local lighting effects that the employed BRDF fitter may not account for, there often is an inherent discrepancy between the simplified normal+BRDF model and the complex optical behaviors of many real-world objects, including layering, where different materials are stacked and interact in complex ways or subsurface scattering, where light penetrates the surface, interacts with the material beneath, and then exits at a different location. These phenomena are challenging to model accurately with standard single normal+BRDF representations. In particular, the shape space of the reflectance function defined by the chosen BRDF model may not be capable of closely matching the input data given the local geometric normal, especially in the common case of relighting using a local illumination model. For this reason, several approaches try to expand the solution by jointly minimizing shape and appearance together in a single framework (e.g., through alternating minimization techniques) [3, 16, 24]. Unfortunately, they increase the system's computational complexity, are less controllable than separate photometric stereo and BRDF fitting solutions, are more difficult to upgrade, and often do not ensure proper convergence [23].

In this paper, we propose a practical approach to the problem that does not break the modularity of classical shape and material characterization steps, but, at the same time, introduces a sort of modeling bridge to strive to enhance the compatibility between geometric characterization and physically-based optical surface modeling. This is done by expanding the solution space by decoupling the normals used for shading from the geometric normals, as done in many rendering frameworks. We consider the normal coming from Photometric Stereo a geometric representation that can be slightly tuned to ensure a better fit of the low-degree-of-freedom BRDF model with the measured data. In particular, we propose to include a new additional normal refinement module in the middle of the shape and material reconstruction pipeline that aims at improving normal (and indirectly the BRDF) computation by imposing a monotonicity constraint based on the common and general assumption that the recovered BRDF should be a monotonic function of the angle between the shading normal and the bisector of lighting and viewing directions (Sec. 3).

For each pixel, the input of our new module is the computed geometric normal and the corresponding measured reflectance sampled across a bi-variate slice (we use the parameterization proposed by Rusinkiewicz [49]). This information is easily obtainable by knowing the pixel appearance profile and the corresponding tuples of normal, view, and light directions. By introducing and utilizing a new metric and optimization framework, we can estimate

the monotonicity level of the sampled reflectance for a candidate set of normals in the neighborhood of the geometric normals, and select the one that performs best according to a specific cost function. Starting with the given geometric normal (from the previous geometric step) as an initial guess, we launch a global optimization process for maximizing the monotonicity metric by finding the optimal pixel shading normal. After that, we use this refined shading normal to fit the dense analytic pixel BRDF. Unlike methods that require coupled shape and material solvers, this approach integrates seamlessly into existing reconstruction pipelines, allowing flexible combinations of Photometric Stereo methods, BRDF models, and BRDF fitters. Users can choose to utilize our corrected normals as improved geometric normals, particularly when using simple normal recovery techniques, or, in the most general case, to use separate normal maps for geometric analysis and visualization. The novel contributions of our work can be summarized as follows:

- A normal refinement module that effectively separates geometric normals from shading normals and is straightforward to integrate into general shape and appearance characterization frameworks. This module operates on a pixelwise basis, taking as input the pixel PS normal and its sampled BRDF. It utilizes a BRDF monotonicity constraint and employs a global optimization framework with BRDF-based domain adaptation. Additionally, it introduces a novel regularization term to prevent normal estimation divergence.;
- A novel BRDF monotonicity metric computed within the domain of the 2D BRDF slice as defined by the Rusinkiewicz [49] parameterization. This metric evaluates BRDF monotonicity with respect to the cosine of the bisector between the view and light directions. It incorporates the duality of diffuse and non-diffuse BRDF signal composition and uses an asymmetric normalization weighted by the gradient magnitude. The metric is designed to be robust against tiny variations and oscillations in the BRDF value caused by measurement noise.

The method's effectiveness is validated through extensive tests on synthetic and real-world controlled datasets, as well as on use cases involving Cultural Heritage assets (Sec. 4).

2 Related work

Acquiring and processing Multi-Light Image Collections involves tasks such as modeling relightable images, computing shape (e.g., in terms of normal maps by Photometric Stereo approaches), and optical surface characterization through Spatially-Varying Bidirectional Reflectance Distribution Functions (SV-BRDF). Providing exhaustive coverage of these topics is beyond the scope of this work, and we direct interested readers to established surveys for a comprehensive overview [18, 28, 39]. In the following, we will discuss the positioning of our methodology within the current literature, motivating its rationale, analyzing current state-of-the-art solutions closely related to our technical contribution, and describing how our practical approach can enhance them directly or from a system-based point of view.

2.1 Relightable images

Relightable images are compact representations that directly map, perpixel, the lighting parameters to the rendered color, eliminating the need for explicitly determining a separation between shape, material, and lighting interaction [42]. When only visual exploration and virtual relighting are required, they often are the prevalent representation. A classical approach in this domain is Polynomial Texture Mapping (PTM) [32], which stores pixel-wise coefficients of a bi-quadratic polynomial representing color variations based on incident light direction. Building on this foundation, subsequent approaches aimed to enhance visual performance through the introduction of different fitting functions [65] or by improving the numerical robustness of the solver [11, 40]. Given that lighting direction and other radiometric properties are often better approximated with functions defined on spherical surfaces, some methods advocate for specific basis functions like Hemi-Spherical Harmonics (HSH) [6, 12]. Other authors, instead, building on the multi-scale nature of the reflectance field, proposed Discrete

Modal Decomposition (DMD), achieving comparable results [43]. Recent approaches leverage neural networks and deep learning to directly learn relightable images from the original Multi-Light Image Collection (MLIC) [44–46, 64]. While learning approaches offer advantages such as the ability to model complex optical behaviors like shadows and interreflections, they come with drawbacks. These include a strong dependence on training data, a high likelihood of producing hallucinations and artifacts, and an inference cost that is generally higher than the cost of evaluating classic analytical formulations.

The compactness and low complexity of relightable images make them highly suitable for fast interactive relighting in both local and remote visualization. For this reason, PTM and HSH are, nowadays, de-facto standard formats in relighting tools from MLIC data, highlighting their widespread adoption mostly in Cultural Heritage application scenarios. Despite their advancements, relightable images are primarily suitable only for qualitative visualizations, with limitations both in high-frequency representation [11] and measurability of object properties. In particular, they are mostly tuned to replicate illumination settings that closely match the measurement conditions. For this reason, they do not provide a physics-based characterization of the surface independent from the measurement illumination. Moreover, the lack of separation between shape and material components makes integrating such representations into common rendering pipelines difficult, for example to insert them into scenes or simulate shape-dependent effects such as shadows (not to mention global illumination). For these reasons, practical solutions to capture separate surface geometry and optical response are the focus of much research [39]. Photometric Stereo and BRDF fitting are standard approaches to address those issues.

2.2 Photometric Stereo

Photometric Stereo estimates surface normals from MLICs. The seminal work was introduced in the 1980's [54, 60]. Conventional Photometric Stereo[1, 53] is based on Lambert's reflectance model, which assumes image intensity proportional to the cosine between the surface normal and the light direction. Unfortunately, all those methods cannot cope well with non-diffuse phenomena such as gloss, highlights, specularity, or shadows. Three main kinds of approaches have been adopted toTo address a non-Lambertian optical response. The first class of methods keeps the Lambertian model and considers non-Lambertian reflectance values as outliers [26, 34, 35, 61, 62]. The second class tries to fit normals with more complex reflectance models that analytically include non-Lambertian observations [15, 21, 48, 52, 57]. The third and last group consists of completely data-driven techniques, and include both example-based approaches [19, 24] and methods that rely on deep neural network architectures [7, 50]. Those techniques can achieve reliable normal map reconstructions in terms of low mean angular errors [27], and they are extremely powerful if the final scope is to exclusively model surface geometry. However, if the normal is used as a preliminary step to compute the optical response of the surface (e.g., BRDF) under local illumination, it must also meet physical constraints to make it compatible with the reflectance model. In particular, many methods try to minimize normal map errors also under some types of reflectance monotonicity constraints [37, 51, 53]. Higo et al. [20] rely on the monotonicity with respect to the cosine of surface normal and light direction to model many non-Lambertian diffuse reflectances. Many other works instead focus on specular lobes characterization, which is best addressed by considering monotonicity with respect to the cosine of the bisector between the view and light directions, and the surface normal [3, 20, 37, 51]. In these works the monotonicity metrics are computed simply by using the signum-based method [17, 33], which locally checks the sign of the derivatives, and computes its normalized sum. Values equal to ± 1 mean pure monotonicity, while a completely non-monotonic function has a value of zero. We also employ a monotonicity constraints but in the context of a refinement step that improves SV-BRDF fitting starting from geometric normals rather than as a means to compute normals from measurements.

2.3 SV-BRDF fitting

A large class of objects can be represented by augmenting normal maps with per-pixel Spatially-Varying Bidirectional Reflectance Distribution Functions (SV-BRDF). A sub-class of methods that extract optical surface behavior in terms of SV-BRDF [18] can accept MLIC data as input signal. This is a specific condition sincea MLIC represents only a slice of the bigger multi-dimensional BRDF space. Moreover, for practical reasons, this slice is typically sparsely sampled, which makes the problem even harder and more ill-posed. Several strategies have been implemented to overcome those issues. Some methods increase sampling by adding more viewpoints and by using additional instruments for normal map bootstrapping [63]. Single-viewpoint approaches add either some particular assumptions to the BRDF nature (e.g., single non-spatially varying BRDF), or try to augment the sampling for each surface position. For instance, Aittala et al. [2] leverages a constant normal map condition (known flat geometry) to extract SV-BRDFs. Providing knowledge or some kinds of assumptions related to the shape [5, 66] and/or the lighting [47] are standard ways to facilitate the system for material extraction and to make the computation more robust. Unfortunately, those methods can not be generally applied to surfaces with complex shapes and a heterogeneous set of different materials. A strategy to both deal with sparse sampling and to bound physically plausible material representations is to rely on a dictionary-based fitting framework [23, 24, 29, 56]. Besides specific implementations, the general rationale behind this sparse fitting and final regularization is that the BRDF at each pixel is represented as a weighted combination of reference BRDFs (or the non-negative span of a dictionary). Similarly, Tingdahl et al. [55] build a dictionary of base materials by directly extracting them from a data-driven clustering step. Similarly to relightable images (Sec. 2.1) and Photometric Stereo (Sec. 2.2), SV-BRDFs can be modeled by neural networks, and a lot of recent works exploit the representation power of such networks to learn and characterize the surface optical behavior from dense or sparse data[8, 14, 30]. The vast majority of those techniques compute independent per-pixel BRDFs, while other techniques try to extract pixel material by looking at its neighborhood, hoping to increase BRDF sampling under the assumption that nearest pixels exhibit similar (or same) optical responses but with different normals [38].

In general, these methods strongly rely on the previous geometric step, i.e., the estimation of the pixel-wise surface normal; proper knowledge of normals is extremely important mostly for materials that exhibit a highly specular behavior. However, this is a weak assumption, since MLIC data of non-Lambertian materials often affect Photometric Stereo algorithms and they produce significant deviation from the correct solution in terms of normal angular error and capability to reproduce the measured appearance using a local normal+BRDF model; for instance, as already mentioned in Sec. 2.2, sometimes they violate the monotonicity constraint. In those cases, a proper characterization of the material is affected by a quite deformed BRDF sample distribution; sometimes a specular or metallic material might be erroneously recognized as diffuse. For this reason, some methods try to iterate an alternating estimation of normals and BRDFs [9], by using the previous normal (or BRDF) characterization as an initial guess for the next step of BRDF (or normal) calculation. Although powerful, the downside of alternating minimization approaches lies in their computational expense and a significant reliance on obtaining a favorable initial solution. This challenge arises due to the non-convex nature of the underlying problem, which is fraught with local minima [23]; this is typically true when the normal error is high and the sampling of the specular high-frequency BRDF region is scarce. Moreover, it should be noted that, in real conditions, local illumination is strongly affected by non-local illumination effects (e.g., inter-reflections), making this optimization even harder. Most solutions use local minimizers, and there is no guarantee that the starting solution is in the basin of convergence, especially when there is only very sparse sampling, as in MLIC conditions.

2.4 Our contribution

The preliminary considerations made in Sec. 2.1 about the main dilemma about implicit or explicit computation of shape and material, together with the limitations of relightable images with respect to a more quantitative application scenario and an easier and more flexible integration into standard rendering pipelines, serve as the primary motivation behind the presented method, which focuses and relies on explicitly modeling the shape (i.e., normal maps) and optical features (i.e., SV-BRDF) of a surface. Our method proposes an approach to enhance the computation and separation between these components and their proper characterization; this improvement is achieved through a technique that includes a novel way to impose a monotonicity constraint in the estimation of the surface normal map, similar to that class of works discussed in Sec. 2.2. Compared to the state-of-the-art methods, which integrate the monotonicity constraint with a straightforward metric within the Photometric Stereo computational framework, we propose here two main pieces of novelty.

First of all, we adopt a highly modular shading normal refinement strategy, which can be easily integrated into a general framework of shape and material characterization. We do not force our monotonicity metric and normal optimization to be used with particular Photometric Stereo methods, SV-BRDF fitting strategies, or combined approaches (see Sec. 2.3). Conversely, we moved toward a completely separate and standalone normal refinement step. As a result, any existing pipeline that produces a normal map (i.e., a normal for each MLIC pixel) and an SV-BRDF (i.e., a set of reflectance parameters for each MLIC pixel can easily integrate the presented approach to improve the final optical and geometrical surface representation. This makes our method extremely accessible, versatile, and flexible.

Moreover, concerning the specific technique, while current Photometric Stereo methods exploiting monotonicity typically exploit the signum-based metric rationale, we instead incorporate an asymmetric normalization process, which is weighted by the gradient magnitude of the BRDF. While assessing the BRDF monotonicity in relation to the cosine of the angle bisector between the viewing and lighting directions, our metric leverages the dual nature of diffuse and non-diffuse BRDF signal components, and is designed to be robust, effectively handling minor variations and oscillations in BRDF values that might arise from measurement noise. In addition, we design and introduce a novel regularization term that aims at minimizing the deviation between the initial Photometric Stereo normal and the final shading normal, thereby preventing excessive divergence in the refinement process in low-error areas.

Our main purpose is to address and overcome problems due to the presence of two critical coexisting conditions, i.e., the need for robust Photometric Stereo methods to manage outliers, and the extremely sparse sampling of the specular lobe of the BRDF caused by the natively MLIC setting. Small amounts of sparse specular samples placed in the wrong region within the BRDF space (this is due to an error in the former normal computation) are typically overlooked or considered as outliers in more complex, robust multi-variate optimizations (e.g., normal/BRDF alternating minimization). In those cases, the specular behavior is lost and the algorithm converges to a wrong diffuse surface model. Our approach is capable of dealing with such extreme and challenging cases.

3 Method

Fig. 1 illustrates the modular framework upon which our proposed contribution relies. Red boxes are processing modules, while other colored boxes are input or derived data. General standard or state-of-the-art techniques are depicted as background grey boxes and they are not the subjects of our study, while the contribution of this paper resides in the *Normal Refinement* module (background yellow box). The pipeline begins with the input of a Multi-Light Image Collection (MLIC) of an object, accompanied by geometric and radiometric calibration metadata, along with a single mask delineating the region of computation. Initially, the pipeline includes a pre-processing module for computing the surface normal map. This module employs a PS approach to generate initial surface normal estimates from the input MLIC. All input and computed data contribute to the estimation of

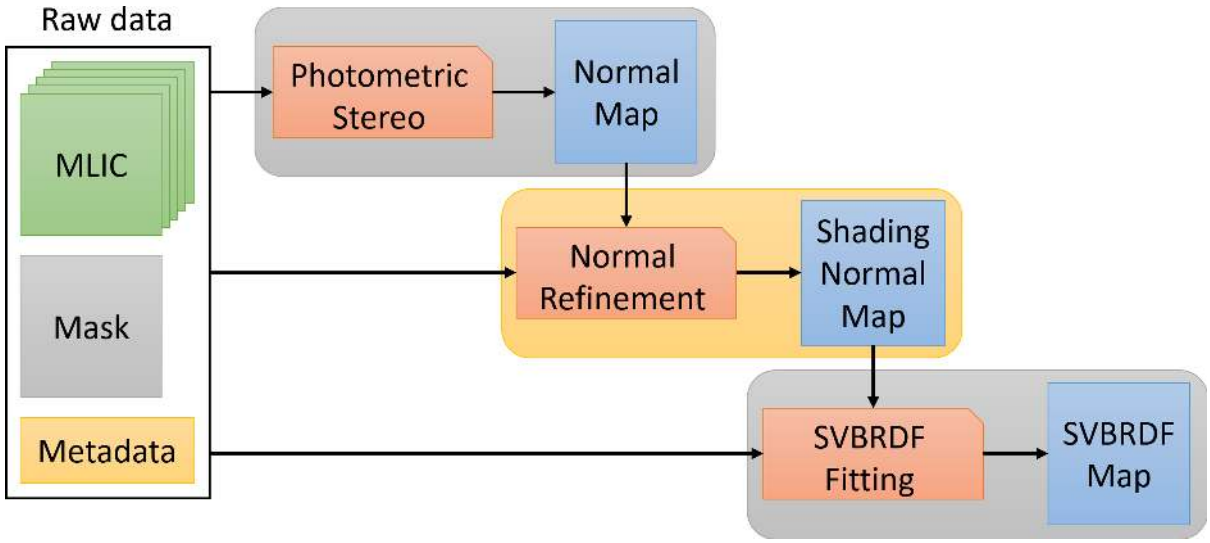


Fig. 1. Overview of the MLIC-based framework on which our proposed method is based. Red boxes are processing modules, while other colored boxes are input or derived data.

refined shading normals, which are then utilized to fit an SV-BRDF model. Our framework lets the Photometric Stereo and SV-BRDF fitting modules be as general as possible, and any traditional or advanced method can be employed. Further details on implementation choices made for these modules in our specific case are elucidated in Sec. 3.2.

3.1 Normal refinement

The proposed normal refinement strategy is based on three main pillars, i.e., the choice of a proper generic BRDF representation (Sec. 3.1.2), the definition of a monotonicity metric (Sec. 3.1.3), and an optimization routine (Sec. 3.1.4). These three elements are chosen to take into account the sparse nature of the BRDF sampling, the geometric constraints of the capture setup (e.g., fixed camera condition and light directions across a hemisphere), and a common surface radiometric behavior (in our case we consider isotropic materials). Before an in-depth evaluation of our method using synthetic and real-world datasets, and solely to enhance clarity and facilitate visual understanding, we exclusively utilize a straightforward synthetic dataset throughout this section (Sec. 3.1.1).

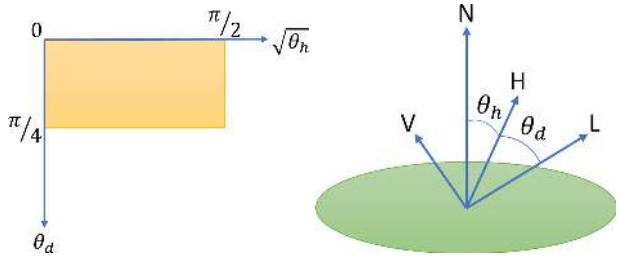
3.1.1 Synthetic Dataset. We generated one simple synthetic dataset by first establishing a material model as our foundation. Specifically, we adopted a dichromatic isotropic Ward model [58], characterized by three primary attributes: diffuse color, specular color, and monochromatic glossiness value. To construct the dataset, we produced the maps corresponding to the surface normals and the attributes. As illustrated in Fig. 2, the dataset consists in a *Hemisphere* positioned on a diffuse Lambertian yellow plane perpendicular to the view direction. The *Hemisphere* exhibits a gray diffuse component. Its left portion demonstrates considerable specularity (featuring a white specular color and a glossiness of 0.5), while the right section represents a perfect Lambertian material (with zero specular and glossiness signal). For the α parameter of the standard isotropic Ward model, we define the glossiness as:

$$g = g = \frac{1 - \alpha^{\frac{1}{e}} - g_{min}}{g_{max} - g_{min}} \tag{1}$$

(a) (b) (c) (d) (e) (f) (g)

Fig. 2. To aid in comprehending the proposed normal refinement algorithm, we present a simple synthetic dataset. This dataset comprises several crafted maps, including the (a) normal, (b) diffuse, (c) specular, and (d) gloss maps. The three images on the right (e), (f), (g) offer glimpses into the simulated Multi-Light Image Collection (MLIC). The simulation employs collinear lighting.

where maximum and minimum gloss respectively are $g_{min} = 1 - \alpha_{max}^{\frac{1}{e}}$ and $g_{max} = 1 - \alpha_{min}^{\frac{1}{e}}$. The minimum and maximum value of α are $\{0.05, 0.6\}$, defined by following the considerations about material behavior from the book of Dorsey et al. [10]. For the sake of representation, we found that $e = 4$ is a good practical choice. Progressing from left to right in Fig. 2, we display all corresponding maps, namely the normal, diffuse, specular, and gloss maps. In the last three columns, we showcase images generated from the simulated MLIC. Our illumination model employs collinear lighting, and we simulate MLIC images by evenly distributing 49 light directions across the hemisphere surrounding the object.



(a) (b)

Fig. 3. The simplified Rusinkiewicz's parameterization. (a) 2D BRDF domain (yellow rectangle) is spanned between the horizontal θ_h and vertical θ_d axes. (b) θ_h and θ_d are obtained by the triplet (N, V, L_i) defined for each pixel i.

3.1.2 BRDF representation. The representation of a measured (analytically unknown) BRDF for a particular surface point can be obtained from a series of input data, a specific choice of a N-dimensional spatial domain, and a radiometric transformation which normalizes all the BRDF values observed under different lighting conditions. In our case, the input data consists of the MLIC reflectance measurements (the point/pixel appearance profile), the corresponding series of light directions L_i and intensities I_i, the point normal N, and a single viewing direction V. For a MLIC acquisition, the camera has a fixed position and viewing direction, so, for a general camera matrix, each pixel has its own viewing vector which remains fixed within its entire appearance profile. We choose Rusinkiewicz's representation [49] to define a 2D domain to organize the BRDF values spatially. Based on the triplet (N, V, L_i) those values are projected on a domain parameterized along two orthogonal axes. The horizontal axis represents the values of the angle θ_h between the surface normal and the half vector $H = \frac{V+L}{\|V+L\|}$, while the vertical axis is the angle θ_d between the half vector H and the viewing V (or the light L) vector (see Fig. 3). Please note that in Fig. 3a, we consider only a limited segment of the BRDF domain. In standard practical

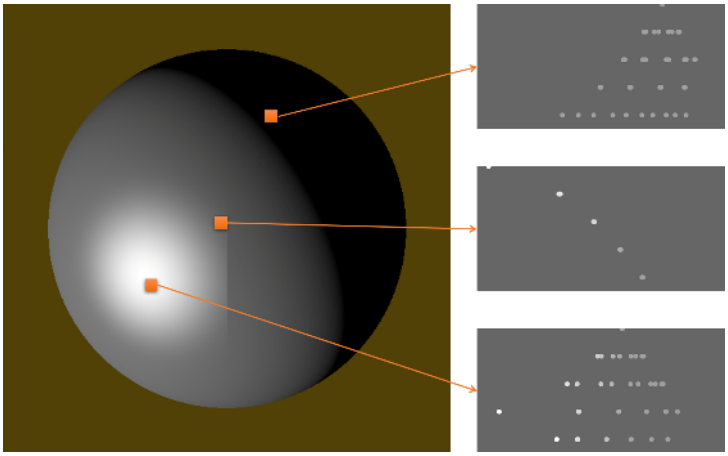


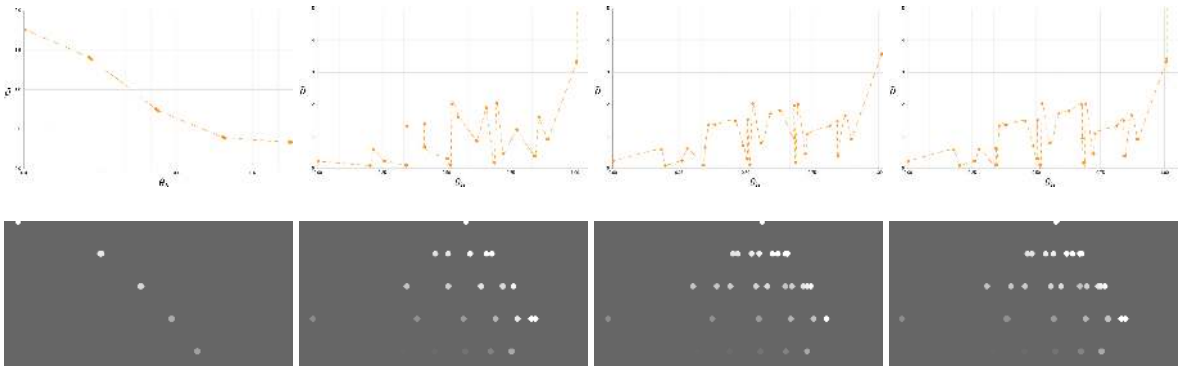
Fig. 4. Sparse sampled BRDFs for three selected points in the *Hemisphere* datasets.

scenarios, during a conventional Multi-Light Image Collection (MLIC) acquisition, the observer views the object from a fixed position while the light source moves across the hemisphere surrounding the object, where the viewer is positioned. Consequently, the angle between the viewing direction V and the light direction L remains consistently below $\pi/2$. This implies a range for θ_d between zero and $\pi/4$, and a range for θ_h equal to $\{0, \frac{\pi}{2}\}$. It is important to emphasize this issue because not only we are dealing with a sparse acquisition, but standard MLICs impose further restrictions on the potential two-dimensional BRDF domain due to the capture practices in Cultural Heritage daily workflows.

Given a general BRDF model $f\,(N, V, L, \dots)$ and a measured reflectance R, we choose the standard cosine-based formulation for the BRDF value normalization, where $R = f\,(N, V, L, \dots)\,(N \cdot L)I$. Each BRDF value associated with a measured reflectance R_i in the appearance profile will be $\rho_i = \frac{R_i}{I_i \cos(\theta_i)}$, where θ_i being the angle between the light direction L_i and the normal N. For a standard (and practical) MLIC, the BRDF representation results in a sparse image as in Fig. 4. In the image, we chose three pixels and reported the corresponding BRDF representations. The first pixel is a completely diffuse one, which results in a series of constant reflectance BRDF samples. For the other two pixels, the difference between a diffuse appearance/region for high θ_h values, and a more specular behavior for smaller θ_h angles is clear. To enhance the visualization of the BRDF sample distribution, we apply the common square root deformation of the horizontal axis (Fig. 3a), which stretches the distribution of the sample along the θ_h dimension.

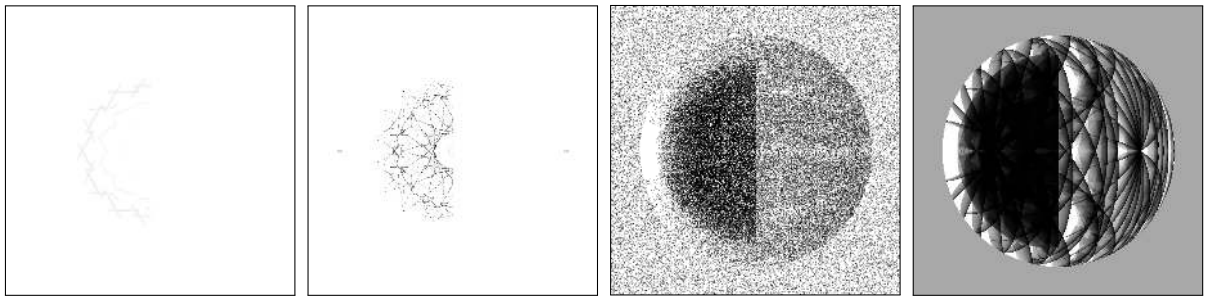
Note that both the spatial organization of the sparse samples and their radiometric values will change with a variation of the surface point normal N; here we consider the view vector V and all light directions L_I and intensities I_i as fixed, given as a set of metadata by a MLIC pre-calibration step. Fig. 5 shows several BRDF representations of the same appearance profile of a specific pixel under different surface normal vectors. We show the sampled BRDF obtained given the correct normal (left-most column), and also, as an example, several other sampled BRDFs obtained by using different normals that all lie on a cone that has the correct normal as axis and an angle $\Delta_\theta = 30°$ as the aperture (this angle also represents the normal angular error); this is useful to see and understand how the BRDF value distribution can vary a lot with the normal, and changes even if the normal angular error remains the same. Considering this aspect, our task here is to find a way to judge how good the given surface normal N might be by analyzing how it influences the spatial and radiometric distribution of measured BRDF samples. We do this by employing a specific metric, which measures how that representation

complies with some chosen physically-based properties exhibited by the BRDF. In our case, we focus on the BRDF monotonicity property. Observe the preservation of monotonicity in Fig. 4 and in the first image of Fig. 5, while it is strongly violated in the remaining BRDF images of Fig. 5. We need to find a way to measure it to assign a score and to choose the proper surface normal.



(a) $\Delta_{\theta,\phi} = \{0, 0\}$, $\mu = 1.0$ (b) $\Delta_{\theta,\phi} = \left\{\frac{\pi}{6}, 0\right\}$, $\mu =$ 2e-4 (c) $\Delta_{\theta,\phi} = \left\{\frac{\pi}{6}, \frac{2\pi}{3}\right\}$, $\mu =$ 7e-3 (d) $\Delta_{\theta,\phi} = \left\{\frac{\pi}{6}, \frac{4\pi}{3}\right\}$, $\mu =$ 6e-5

Fig. 5. Sampled BRDF dependency on normal estimation. We show four possible BRDF images of a point selected from the *Hemisphere* (second row) dataset obtained by considering four different surface normal candidates. The leftmost column shows an image obtained with the correct normal, while the other columns are obtained with normals that all lie on a cone that has the correct normal as axis and an angle $\Delta_\theta = 30°$ as the aperture. The first row shows the dependency of $\tilde{\rho}$ to θ_h, which highlights how the monotonicity is strongly violated with a wrong normal.



(a) Ground Truth (b) Robust Trimmed PS (c) Random error $\Delta_\theta = \left\{0, \frac{\pi}{4}\right\}$ (d) Fixed error $\Delta_\theta = \frac{\pi}{4}$

Fig. 6. Per-pixel Monotonicity Maps for the *Hemisphere* dataset obtained with: (a) the ground truth surface normals; (b) surface normals from Robust Trimmed Lambertian Photometric Stereo; (c) ground truth normals deformed by a random angle drawn in the interval $\Delta_\theta = \left\{0, \frac{\pi}{4}\right\}$; (d) ground truth normals all deformed by an angular error $\Delta_\theta = \frac{\pi}{4}$.

3.1.3 BRDF monotonicity metric. We devised an operator to compute the proposed monotonicity-based metric, focusing on the relationship between the angle θ_h and the normalized BRDF value. Initially, it eliminates the dependency on θ_d by projecting all BRDF samples onto the horizontal axis. Then, it further projects the multi-spectral BRDF value onto the hyper-diagonal of the multi-spectral hyper-cube by computing $\tilde{\rho} = |\bar{\rho}(\theta_h)|$, where $\bar{\rho}$ represents the projected BRDF value and $\tilde{\rho}$ denotes its amplitude. The functions are plotted above the

corresponding BRDF images in Fig. 5.The monotonicity of these functions is evaluated using a metric inspired by the standard *Signum Formula*, yet with several refinements: (1) consideration of the duality of diffuse and non-diffuse signal composition of the BRDF; (2) asymmetric normalization weighted by the gradient magnitude; (3) incorporation of a quick estimation of the albedo value to enhance the metric's robustness to tiny variations and oscillations in the BRDF value due to measurement noise; (4) definition within the interval $\{0, 1\}$, where 1 indicates decreasing monotonicity (typical BRDF behavior), 0.5 represents a constant function (e.g., perfectly Lambertian diffuse case), and 0 signifies perfectly increasing monotonicity.

For each pixel, we begin by setting the initial value of the monotonicity to $\mu = 1$. Then, we analyze the curve defined by a series of data points, specifically the values of $\tilde{\rho}$, by examining each pair of adjacent data points along the curve. When comparing each pair, we calculate the difference between the $\tilde{\rho}$ value of the current point i and that of the next point $i + 1$, referred to as δ. If δ is bigger than a tolerance value, indicating an upward trend between the two points and thus a violation of the expected BRDF's decreasing monotonicity, we update the monotonicity μ by reducing it according to the magnitude of δ. This update is computed as $\mu_{new} = \mu_{old}/(1 + \delta)$. Upon examining all data points in the curve, the last updated (current) value of μ becomes the final monotonicity metric. The tolerance ϵ takes into account both the minimum absolute numerical error ϵ_{abs} in the BRDF values, and a relative tolerance η with respect to the diffuse signal *albedo*. We find that setting $\eta = 0.1$ and ϵ_{abs} = 1e-4 are sufficient to mitigate the effects of noise. The algorithmic schematic in Algorithm 1 outlines the computation process for the monotonicity metric. Here, $\tilde{\mathbf{P}}$ represents the set of values of $\tilde{\rho}$ corresponding to the chosen pixel.

Algorithm 1 Monotonicity Metric Computation

function MonotonicityMetric($\tilde{\mathbf{P}}$, albedo, η)
  $\mu \leftarrow 1.0$
  $\epsilon_{abs} \leftarrow$ 1e-4
  $\epsilon_{rel} \leftarrow (1 + \eta) \times$ albedo
  $\epsilon \leftarrow max\,(\epsilon_{abs}, \epsilon_{rel})$
  $N \leftarrow$ size of $\tilde{\mathbf{P}}$
  for $i \leftarrow 0$ **to** $N - 1$ **do**
    $\delta \leftarrow \tilde{\rho}[i + 1] - \tilde{\rho}[i]$
    if $\delta > \epsilon$ **then**
      $\mu \leftarrow \mu \times \frac{1.0}{(1.0+\delta)}$
    end if
  end for
  return μ
end function

Fig. 5 reports also the value of μ for each BRDF; correct shading normals produce high μ values, while inconsistent shading normals exhibit extremely low monotonicity values (μ closer to zero). In addition, for, Fig. 6 shows four monotonicity maps of the *Hemisphere* dataset. One map is computed with ground truth normals. Another has been obtained by considering a normal map estimated with a robust trimmed PS approach. In the third case, we apply a random angular error to the ground truth normals in the range $\Delta_\theta = \left\{0, \frac{\pi}{4}\right\}$, while in the last four image we globally apply an angular error equal to $\Delta_\theta = \frac{\pi}{4}$.

3.1.4 Shading normal computation. For each pixel in the MLIC, we initialize the shading normal with the normal N computed by the Photometric Stereo approach. Then we build the BRDF representation as explained in Sec. 3.1.2. The shading normal $\tilde{N}$ is then obtained by launching an optimization procedure that maximizes the monotonicity

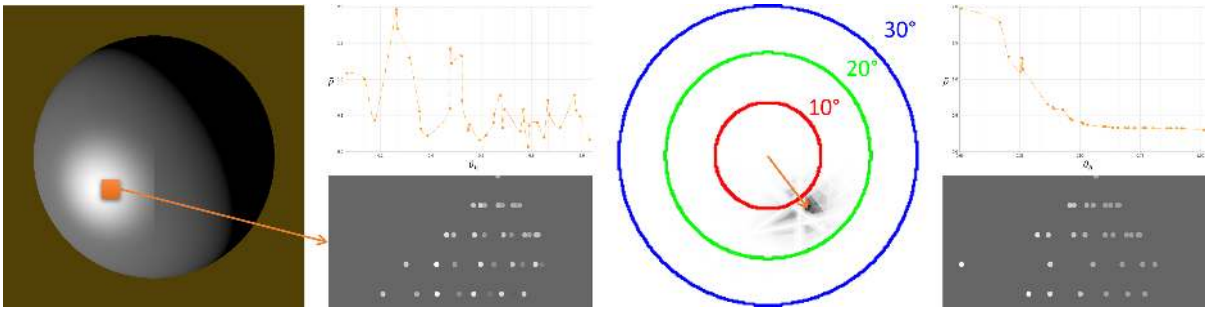
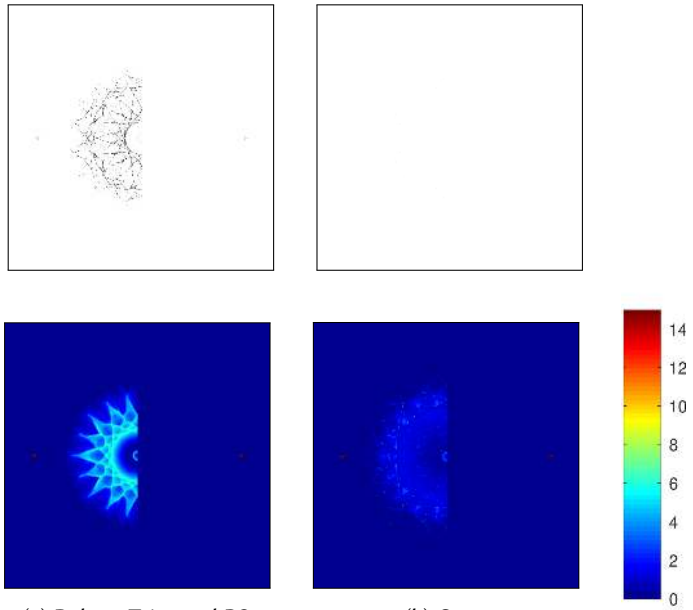


Fig. 7. Shading Normal Computation. A pixel is chosen from the *Hemisphere* dataset. Initially, the coarse normal yields a significantly non-monotonic BRDF signal (second column, μ = 8e-5, angle error $\epsilon = 15°$). The minimum of the cost function (third column) determines the correction (orange arrow) we need to apply to the coarse normal in order to obtain the refined normal. This refinement results in a monotonic behavior of the sparsely sampled BRDF (fourth column, $\mu = 0.99$, angle error $\epsilon = 1.23°$).



(a) Robust Trimmed PS (b) Ours

Fig. 8. Monotonic maps (first row) and the corresponding normal angular errors (second row) computed between the ground truth normal maps and those obtained by (a) the *Robust Trimmed PS* method and (b) the proposed refinement approach.

as

$$\tilde{N} = argmin_N \left[1 - \left(\mu\left(N\right) + \mathbf{R}\left(N\right)\right)\right] \tag{2}$$

where $\mathbf{R}(N)$ denotes a regularization term aiming to minimize the deviation between the initial normal and the final one, thereby preferring the original photometric normal when the error is low. As we initiate the minimization within a canonical search space where the normal initial guess aligns with the z-axis, the regularization term simplifies to $\mathbf{R}(N) = \lambda N_z$. We found that a small value of λ = 1e-4 works well in all our tests.

Given the high sparsity of the measurements, the fact that only a small number of those BRDF samples are in the numerically significant region (most of them lie in the constant diffuse zone), and the response of the monotonicity field to changes in the normal orientation, we observed that traditional non-linear gradient-based optimization (e.g., the Levenberg-Marquardt algorithm [22, 31]) performs poorly and frequently becomes trapped in local minima. Revisit the signal of the cost function in Fig. 7 or Fig. 9. The presence of numerous local minima, encircled by regions of maximum values in the cost function, is strikingly evident. For this reason, we adopted a bounded, global, brute-force search. In practice, we calculate all monotonicity values (or cost function values) for a dense set of candidate normals within a range of angles equal to or less than a threshold compared to the coarse normal. We choose here a threshold of $\pi/6$.

Fig. 7 depicts the refinement procedure applied to a pixel in the *Hemisphere* dataset. For that particular pixel, we took the ground truth normal and applied a deformation equal to $\Delta_{\theta,\phi} = \{15^\circ, 135^\circ\}$. to apply this deformation, imagine that we anchor a canonical frame to the normal, where the z-axis is collinear to the normal; Δ_θ is an angle with that z-axis while Δ_ϕ is a rotational angle around that z-axis. We consider that deformed normal as the initial surface normal computed by the Photometric Stereo step, and we want to refine it to obtain the ground truth normal. The second column of Fig. 7 shows the corresponding BRDF image and the plot of the function $\tilde{\rho}(\theta_h)$. We can see how the normal deformation deteriorates the monotonicity of the function, with a computed value of μ = 8e-5. The third column shows the cost function, which represents a 2D search space domain for the refined shading normal. In particular, the initial normal is located at the center of the domain and the circular regions correspond to cones around it; more specifically, red, green, and blue circles are the cones of angle 10, 20, and 30 degrees. We display only a segment of the normal search space to enhance visual clarity. We could of course extend the search to angles greater than 30 degrees. The minimum of the cost function, indicated by the orange arrow, corresponds with the correction we need to apply to the Photometric Stereo normal to obtain the optimized shading normal; note how this is the opposite of the deformation we had applied. In the example of Fig. 7 the refined normal has an angular error of about one degree with the ground truth normal. The last column shows how the resulting BRDF image and the plot of the function $\tilde{\rho}(\theta_h)$ are more compliant with the monotonicity constraint. The monotonicity value now becomes $\mu = 0.99$. Finally Fig. 8 shows the angular error of the entire normal field for the *Hemisphere* dataset with respect to the ground truth normal map. We compare the normals obtained with the *Robust Trimmed PS* and those refined with the proposed approach. We also present the corresponding monotonic maps. The color bar depicts the angular error between 0° and 15°. Our approach increases the quality of the reconstructed normal, which globally has less angular error and exhibits a higher monotonicity behavior.

3.2 Implementation

The proposed method is independent of the specific approaches used for photometric normal estimation and BRDF fitting. To implement the entire surface characterization pipeline for presenting results in this article (see Fig. 1), in addition to the proposed normal refinement method, and without loss of generality, we had to choose two default approaches for normal estimation and SV-BRDF computation. By following the survey and the reported statistics of several methods from the work of Shi et al. [53], Wang et al. [59], and for the sake of practicality and the general applicability of our modular framework, we selected, as already mentioned in the Sec. 3.1.4 and Fig. 8, the *Robust Trimmed PS* to compute the surface normals. To characterize the surface optical response in terms of SV-BRDF, we adopted the formulation and fitting framework proposed by Pintus

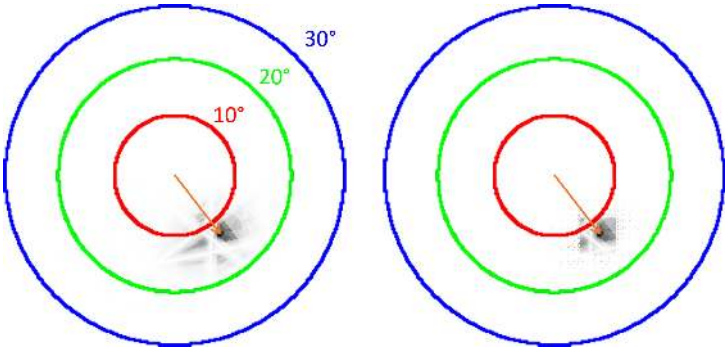


Fig. 9. Comparison of a dense cost function sampling with a sparse one achieved through sparse-to-dense multi-level optimization, which progressively focuses on the most promising region until determining the optimal normal.

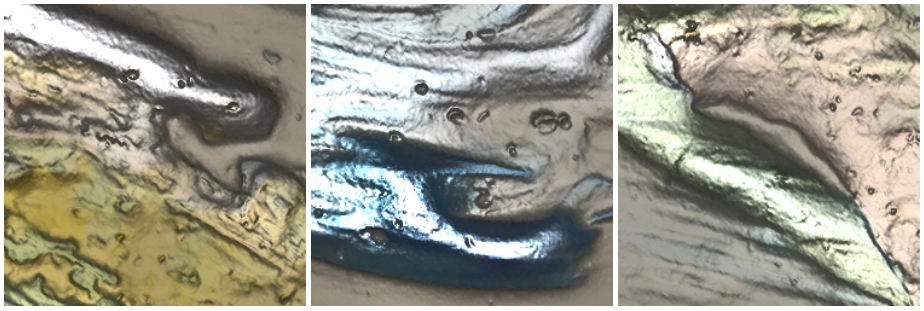
et al. [38]. Sec. 4.1 illustrates results where we mix and match different normal computation components and BRDF representations.

Photometric Stereo angular errors and problems due to normals that do not respect the monotonicity constraint typically concern surface points that are not purely diffuse. To speed up the computation, besides exploiting the embarrassingly parallel nature of the pixel-wise computation, we also use conservative heuristics to avoid the optimization of perfectly Lambertian BRDFs. This is done by performing the optimization only for the pixels whose BRDF representations significantly deviates from the constant albedo value BRDF representations that deviate too much with respect to their albedo value. In practice, we consider a pixel as diffuse if $max\,(\tilde{\rho}) - k_d > k_d - min\,(\tilde{\rho})$, where k_d is the pixel albedo value computed by the previous SV-BRDF fitting step, and $max\,(\tilde{\rho})$ and $min\,(\tilde{\rho})$ are the highest and lowest projected BRDF values. For diffuse pixels we keep the normal computed by the Photometric Stereo approach.

Furthermore, while bounded, a pure brute force computation may result in numerous unnecessary calculations depending on the number of candidate normals. To address this, we employ an iterative sparse-to-dense multi-level approach for globally searching the optimal normal. We begin by defining an angular step that is sufficiently dense to explore the bounded normal space yet coarse enough to reduce the computational cost for each search space span. In each iteration, we progressively decrease the angular step , centering the search region around the current best normal. Fig. 9 presents a comparison between the dense cost function (the same as in Fig. 7), computed for all possible monotonicity values across candidate normals, and the sparse cost function obtained through the iterative procedure. The iterative process begins by calculating monotonicity values for a sparse, uniformly distributed set of candidate normals. It then gradually hones in on the most promising region until the optimal normal is determined.

4 Results

The newly proposed normal map refinement module has been implemented in C++ and seamlessly integrated into our pipeline, as detailed earlier in Sec. 3.2. In this section, we validate the refinement module by conducting a comparative analysis of the pipeline's performance with and without the addition of the normal map refinement. Our primary objective is to demonstrate the enhancement in the final relighting in terms of both numerical accuracy, i.e., Peak Signal-to-Noise Ratio (PSNR), and perceptual quality, i.e., Structural Similarity Index Measure (SSIM) and ꟻLIP metrics. We evaluate our algorithm's efficacy in improving the final reconstruction and relighting quality, particularly for challenging flat objects and visually/geometrically intricate subjects. Before conducting an extensive analysis of our method on multiple synthetic and real-world cases, we first evaluate the normal



(a) Paint-Texture-00 (b) Paint-Texture-01 (c) Paint-Texture-02

Fig. 10. One image is presented from each of the three rendered synthetic datasets.

refinement technique by integrating it into two distinct pipeline instantiations, employing different Photometric Stereo and SV-BRDF fitting models, and running those pipelines on a real dataset (refer to Sec. 4.1). This initial test aims to demonstrate that, due to the modular design of our approach, our method is agnostic to the other techniques integrated into the various modules of the surface characterization pipeline. Then, to investigate and evaluate the enhancements in normal map reconstruction and relighting accuracy, we conduct comprehensive quantitative analyses using synthetic data with ground truth information and painting mockups captured under controlled conditions (see Sec. 4.2). Finally, we present practical results obtained from the free-form acquisition of a pair of cultural heritage items, specifically paintings (refer to Sec. 4.3).

Throughout our tests, we consistently employ a small set of metrics. For comparing normal maps, we utilize a straightforward angular error measurement. Summarizing the quality of relighting involves plotting graphs depicting the relative percentage improvement in both *PSNR* and *SSIM*. Denoting m_0 as the metric value (*PSNR* or *SSIM*) representing the error between the original and the relighted image obtained with the non-refined normal, and m_1 as the metric computed with the refined normal, the percentage relative improvement is defined as follows:

$$\Gamma = 100\frac{|m_1 - m_0|}{m_0}\% \tag{3}$$

The *PSNR* and the more perceptually-oriented *SSIM* results are complemented by incorporating the ꟻLIP metric [4]. This metric serves as a difference evaluator, effectively approximating the distinctions perceived by humans between a rendered image and its corresponding ground truth image. The central concept behind ꟻLIP is to present the perceptual differences observed by humans under the specific condition of alternating between two perfectly superimposed images without any blank intervals. The metric is therefore designed to filter out overall image disparities that, regardless of their numerical magnitude, cannot be discerned by humans, while emphasizing differences in very localized image regions with point-like shapes or edges, rather than changes in color intensity and chromaticity. Given the extensive array of tests conducted, we showcase the ꟻLIP maps and associated statistics for selected challenging images from the original and relighted MLICs.

4.1 Modularity test

Although the following sections present results using a pipeline that employs *Robust Trimmed PS* as the Photometric Stereo method and the *Ward* optical model as the BRDF fitting formulation, here we would like to demonstrate the flexibility of our approach by showcasing a different test case with two alternative configurations. We want to show that we achieve similar and consistent performance by replacing either the Photometric Stereo method

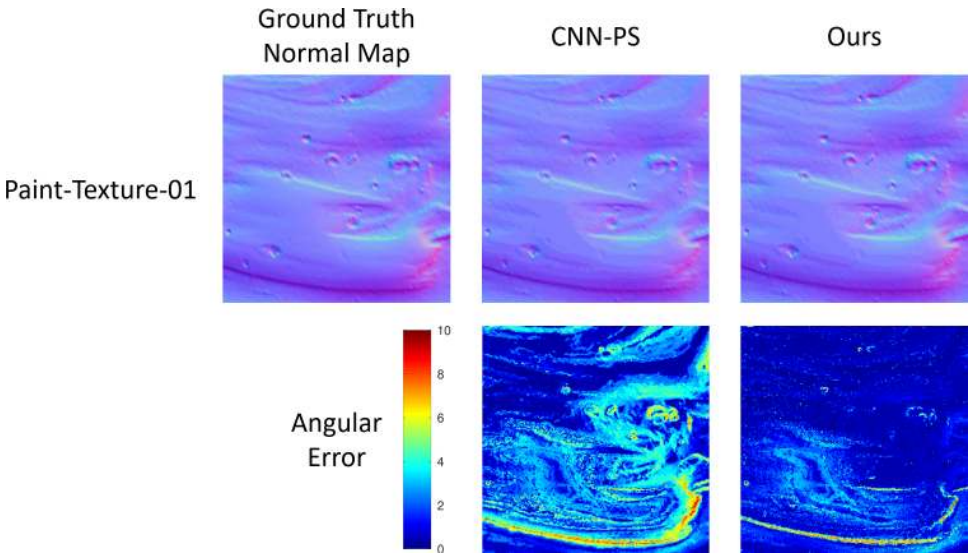
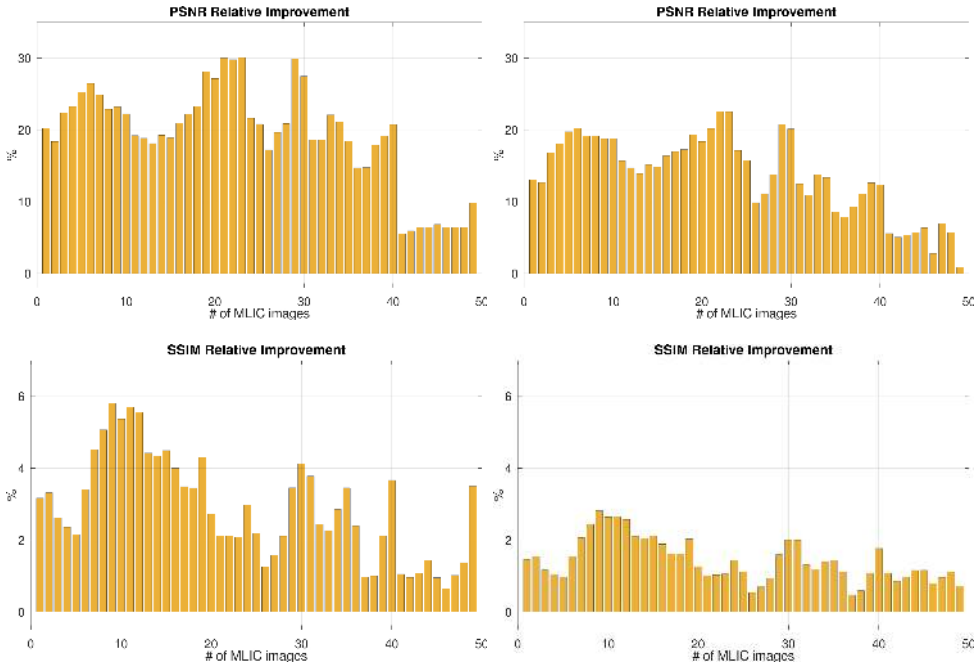


Fig. 11. Angular deviation in the *Paint-Texture-01* dataset between the ground truth normal maps and those computed using both the *CNN-PS* and our approach.

Synthetic Dataset	Method	ꟻLIP Statistics			
		Mean	Weighted Median	1st Quartile	3rd Quartile
Paint-Texture-01	*CNN-PS + Ward*	0.11	0.15	0.09	0.28
	Ours	**0.05**	**0.09**	**0.05**	**0.16**
	CNN-PS + Blinn-Phong	0.15	0.19	0.13	0.31
	Ours	**0.11**	**0.16**	**0.10**	**0.21**

Table 1. ꟻLIP statistics for one image of the *Paint-Texture-01* dataset as the result of two pipelines that utilize *CNN-PS* as the Photometric Stereo method, and either *Ward* or *Blinn-Phong* formulation for the BRDF fitting step.

or both the Photometric Stereo and the BRDF model within our pipeline. This choice illustrates the modularity of our refinement strategy, which is designed to be agnostic to the other modules in the surface reconstruction framework. For the Photometric Stereo method, we select the approach from Ikehata's paper [25], known as *CNN-PS*, i.e., Convolutional Neural Network-based Photometric Stereo. This method leverages the power of deep learning to improve the robustness and accuracy of surface normal estimation and has demonstrated strong performance on the standard *DiLiGenT* benchmark [53]. As an alternative BRDF model, we employ the stretched Blinn-Phong model [36]/REVISIONADD, a simple, physically plausible, but not physically based reflectance model for diffuse and specular materials. Fig. 11 shows the improvement in angular error compared to the normal map computed by *CNN-PS*. This enhancement is also reflected in the relighting statistics, where all relative improvements in *PSNR* and *SSIM* are positive (see Fig. 12). Additionally, Fig. 13 and Tab. 1 report the improvements in ꟻLIP statistics for a relighted image in the *Paint-Texture-01* dataset. This test provides insight into how our method can improve relighting quality when seamlessly integrated into general shape and appearance characterization frameworks.



(a) *CNN-PS + Ward* (b) *CNN-PS + Blinn-Phong*

Fig. 12. The performance of all relit images for the *Paint-Texture-01* dataset is assessed. We present the percentage relative improvement for both the *PSNR* (Peak Signal-to-Noise Ratio) in the first row and the *SSIM* (Structural Similarity Index) in the second row. We present the results for two pipelines that utilize *CNN-PS* as the Photometric Stereo method, and either *Ward* or *Blinn-Phong* formulation for the BRDF fitting step.

4.2 In-depth quantitative evaluation

Our quantitative evaluation is first performed on synthetic tests, where ground truth is available (Sec. 4.2.1) and then on laboratory measurements of painting mockups (Sec. 4.2.2, where acquisition is performed under very controlled conditions.

4.2.1 Synthetic tests. Synthetic tests conducted on rendered models from manually defined maps support the evaluation of the methods in a fully controlled environment where ground truth is available for both normals, BRDFs, and acquisition parameters. We selected a synthetic model from the *EveryTexture* database [13], as it closely resembles the detailed shape and appearance of the real-world objects we are interested in. Utilizing the provided *Diffuse*, *Bump*, and *Normal maps*, we retained the *Diffuse* map as is, generated the *Gloss* component by converting the *Bump* map to monochrome and rescaling it, and assigned a constant highlight color. These maps were then cropped into several regions to create three synthetic MLICs using a fixed camera and 52 directional lights. Fig. 10 displays one of the 52 rendered images for each synthetic dataset. With access to ground truth normal maps for these datasets, our initial test compares the normals computed by a chosen standard method (i.e., *Robust Trimmed PS*) with those obtained using our refinement strategy. Fig. 14 illustrates how, in this controlled case, our approach improves the normal field by significantly reducing the angular error with respect to ground truth normals. In the problematic areas, we observe a gain in normal accuracy of up to 5 degrees. Additionally, we provide error statistics for the virtual relightings obtained with and without our refinement step (see Fig. 15).

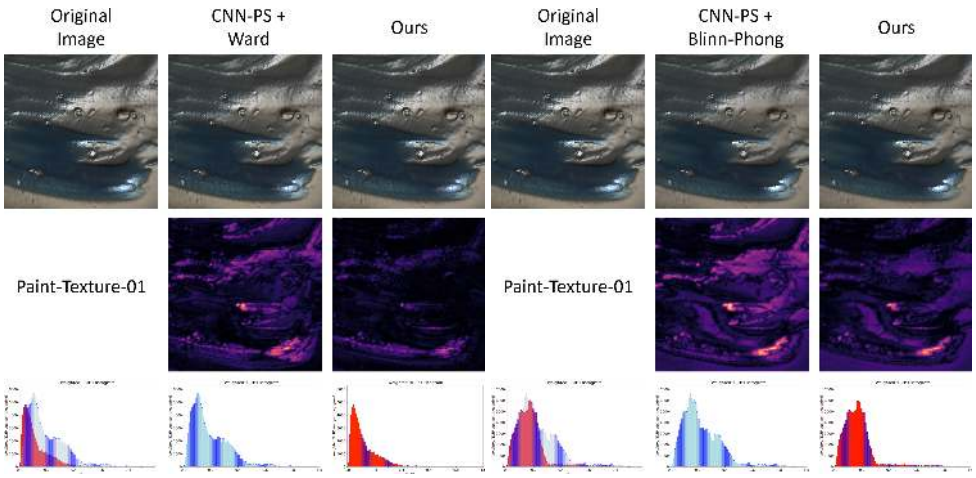


Fig. 13. We juxtapose selected original images of the *Paint-Texture-01* dataset with their digitally relit versions generated from the SV-BRDF calculated using a normal map from a neural network based method *CNN-PS* and our refined normal map. We present the results for two pipelines that utilize *CNN-PS* as the Photometric Stereo method, and either *Ward* or *Blinn-Phong* formulation for the BRDF fitting step. The final two rows showcase the respective ꟻLIP maps and associated statistics. The leftmost plot in the third row of each group of images overlays the two sets of statistics to highlight the improvement more effectively.

Using our two percentage relative improvement metrics, we plot the improvement for all 52 images in the MLIC. Practically for all images, the improvement is positive in terms of both *PSNR* and *SSIM*. Tab. 2 and Tab. 3 summarize the statistics, related to the synthetic tests, of *PSNR* and *SSIM*, respectively. These tables include the minimum, first quartile, average, median, third quartile, and maximum values, with the best results highlighted in bold. Finally, Fig. 16 presents the corresponding ꟻLIP maps and statistics of selected original images from the three synthetic datasets. In the bottom row, the leftmost plot in the third row of each group of images overlays the two sets of statistics to highlight the improvement more effectively. Tab. 4 numerically presents the ꟻLIP statistics in terms of mean, weighted median, and 1st and 3rd quartile of the ꟻLIP values. The best values, highlighted in bold, all belong to the proposed approach.

4.2.2 Real-world painting mockups. In this section, we address the evaluation of our solution within a typical real-world laboratory setting, which can be controlled to ensure the precision and reliability of the statistics and the corresponding evaluation. Our focus centers on six painting mockups, each exhibiting a diverse array of spatially varying material distributions across the surface (refer to Fig. 17). These mockups, originally designed for testing BRDF capture methods [38], were executed on painting paper, employing standard acrylic colors. Our objective was to achieve a wide spectrum of color mixtures and geometries, thereby encompassing different painting styles and effects. In terms of color mixtures, our experimentation spanned from vividly blended hues to distinctly segregated color layers. As for geometry, we explored different configurations, ranging from thinly applied, flat color layers to the rich textures imparted by a heavily laden brush. We pushed the boundaries further, experimenting with pronounced geometric features achieved by directly depositing color from the tube. To enhance the diversity of our mockups, we applied a thin coating of gloss varnish to half of each surface once all color layers had thoroughly dried. This dual treatment aimed to create a juxtaposition between a subtly diffusive surface and an intensely reflective one, thus enriching our evaluation with contrasting material properties.

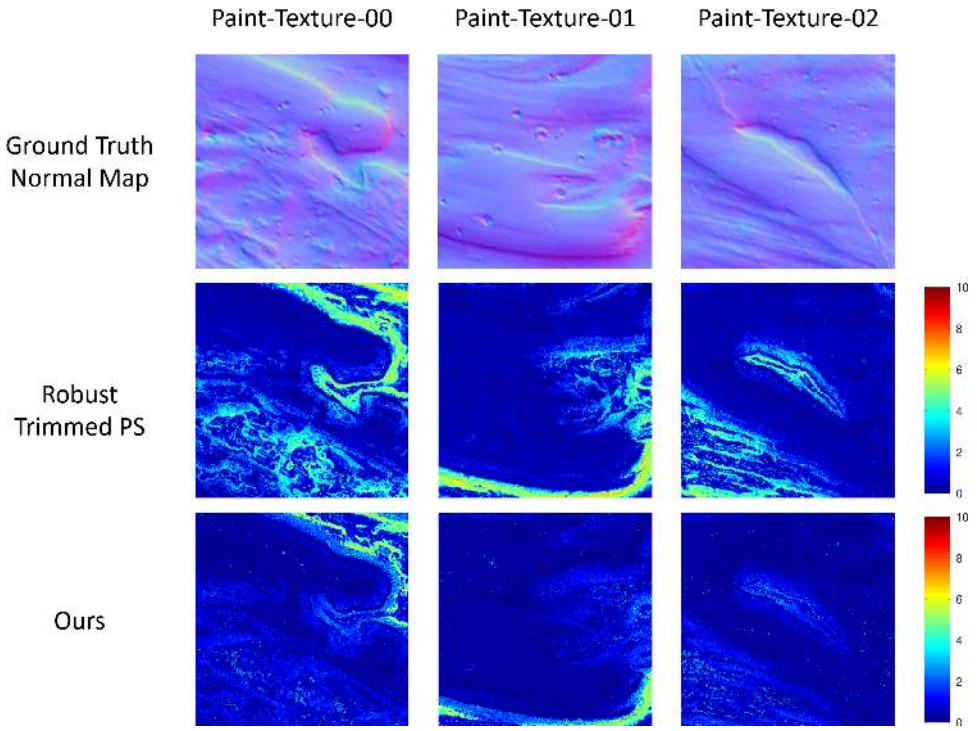
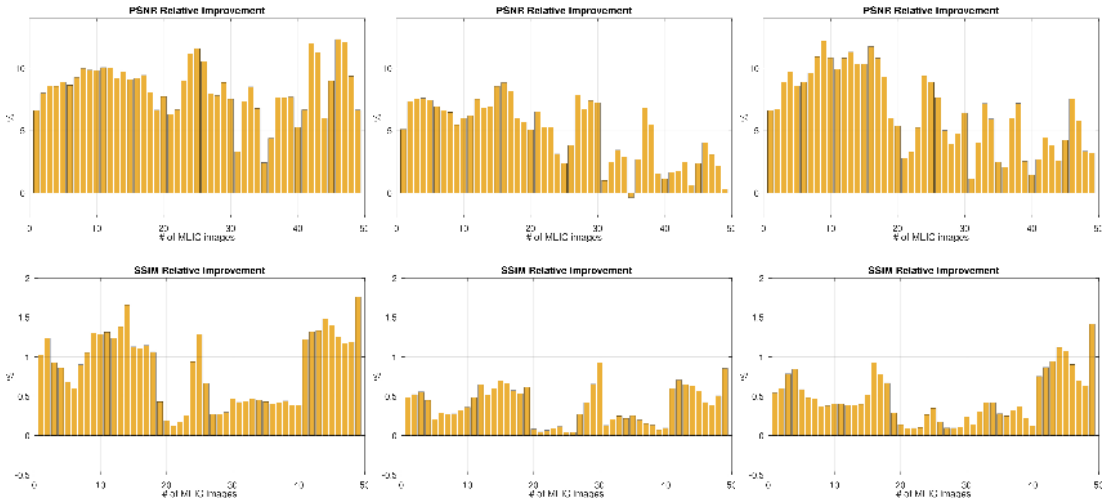


Fig. 14. Angular deviation between the ground truth normal maps and those computed using both the *Robust Trimmed PS* and our approach.

For the acquisition of MLICs for our mockups, we employed a custom-designed light dome with a radius of approximately 30cm. This dome was equipped with 52 neutral white LED lights, ensuring comprehensive coverage across the visible spectrum. Images were taken with a DSLR FX Nikon D810 Camera with 36.3 megapixels with a 50mm AF Nikkor Lens. Four glossy spheres were utilized to recover the light direction, while a *Spectralon* target was captured for flat field light intensity calibration. These calibration techniques were instrumental in guaranteeing consistency, precision, and highly controlled operations throughout the acquisition and processing steps.

Our analysis of error statistics for virtual relightings, comparing results obtained with and without our refinement step, reveals more pronounced benefits of our shading normal computation in this real-world use case than in the synthetic tests (refer to Fig. 18 and Fig. 19). We observed a relative improvement of up to approximately 10% in both *PSNR* and *SSIM*, with notably superior results in images illuminated frontally (higher image numbers) compared to those under raking light conditions (lower image numbers), where shadows and diffuse surface signals predominate. This behavioris consistent with our expectations, as elucidated in Sec. 3, where we discussed the enhanced presence and clarity of monotonicity in surfaces exhibiting highly glossy or specular profiles rather than diffuse surface points. Nevertheless, it is reassuring to note that our algorithm also effectively preserves the integrity of those diffuse appearances, as evidenced by its performance on images captured under raking light conditions. In Tab. 2 and Tab. 3, we respectively report the statistics for *PSNR* and *SSIM* for the real-world painting mockups. The tables provide the minimum, first quartile, average, median, third quartile, and maximum values, with the top-performing results highlighted in bold.



(a) *Paint-Texture-00* (b) *Paint-Texture-01* (c) *Paint-Texture-02*

Fig. 15. The performance of all relit images for the three synthetic datasets simulating flat painting regions is assessed. We present the percentage relative improvement for both the *PSNR* (Peak Signal-to-Noise Ratio) in the first row and the *SSIM* (Structural Similarity Index) in the second row.

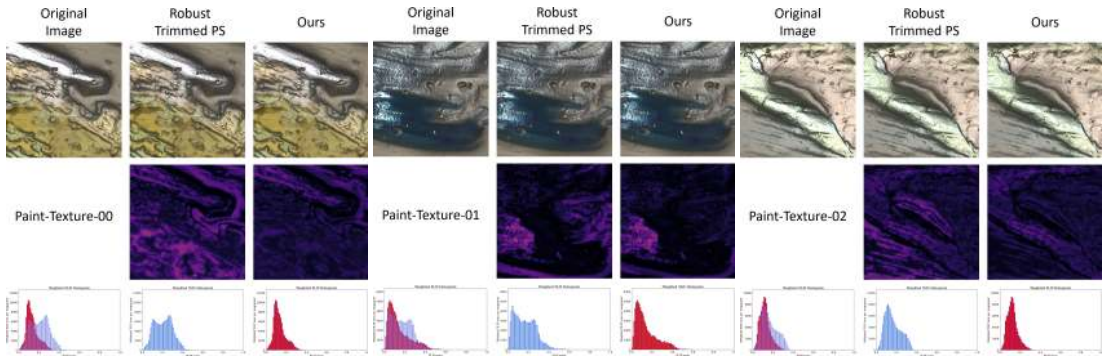


Fig. 16. We juxtapose selected original images of the three synthetic datasets simulating flat paintings regions with their digitally relit versions generated from the SV-BRDF calculated using a normal map from a standard *Robust Trimmed PS* and our refined normal map. The final two rows showcase the respective ꟻLIP maps and associated statistics. The leftmost plot in the third row of each group of images overlays the two sets of statistics to highlight the improvement more effectively.

Concluding our analysis, Fig. 20 and Fig. 21 showcase the corresponding ꟻLIP maps and statistics derived from selected original images within the real-world painting mockups. Within each group of images, the leftmost plot in the third row juxtaposes the two sets of statistics, effectively highlighting the improvement in relighting quality obtained by our approach. For a comprehensive numerical overview, Tab. 5 provides ꟻLIP statistics, including mean, weighted median, and the 1st and 3rd quartiles of ꟻLIP values. Notably, the most favorable values, denoted in bold, are consistently attributed to our proposed approach.

Dataset		Min	Q1	Avg	Median	Q3	Max
Paint-Texture-00	*Robust Trimmed PS*	28.02	36.76	37.69	38.42	39.77	42.88
	Ours	**29.88**	**38.74**	**40.85**	**41.70**	**43.53**	**46.11**
Paint-Texture-01	*Robust Trimmed PS*	22.50	36.70	38.50	39.88	42.22	46.58
	Ours	**22.56**	**37.54**	**40.46**	**42.50**	**45.01**	**48.37**
Paint-Texture-02	*Robust Trimmed PS*	27.82	37.83	39.92	40.62	43.66	47.11
	Ours	**28.71**	**38.63**	**42.62**	**43.93**	**47.46**	**49.40**
Mockup #0	*Robust Trimmed PS*	20.06	30.50	32.56	32.48	35.55	**41.42**
	Ours	**21.04**	**31.51**	**33.03**	**33.17**	**35.73**	41.24
Mockup #1	*Robust Trimmed PS*	19.98	30.80	32.30	33.00	34.77	39.39
	Ours	**21.71**	**31.40**	**33.07**	**33.52**	**35.02**	**39.46**
Mockup #2	*Robust Trimmed PS*	21.62	31.53	33.72	34.63	36.23	**43.30**
	Ours	**22.18**	**32.43**	**34.26**	**34.97**	**36.92**	43.23
Mockup #3	*Robust Trimmed PS*	19.83	33.85	38.28	38.66	42.80	**53.11**
	Ours	**20.95**	**34.96**	**39.22**	**39.87**	**43.74**	52.80
Mockup #4	*Robust Trimmed PS*	28.26	37.15	39.92	39.85	43.82	**51.09**
	Ours	**29.48**	**38.63**	**40.83**	**40.89**	**44.29**	50.89
Mockup #5	*Robust Trimmed PS*	22.65	36.32	37.73	38.80	40.80	43.63
	Ours	**23.73**	**36.55**	**38.13**	**38.95**	**41.08**	**43.94**
Retablo Crop #0	*Robust Trimmed PS*	27.54	32.31	39.22	39.32	45.74	53.25
	Ours	**30.60**	**33.82**	**40.55**	**40.34**	**46.58**	**55.14**
Retablo Crop #1	*Robust Trimmed PS*	20.76	27.49	38.92	39.80	49.55	54.96
	Ours	**24.26**	**29.73**	**40.67**	**41.03**	**50.26**	**56.95**
Retablo Crop #2	*Robust Trimmed PS*	25.18	32.99	46.02	47.77	57.56	62.97
	Ours	**31.15**	**36.71**	**48.43**	**50.55**	**58.94**	**63.23**

Table 2. *PSNR* (Peak Signal-to-Noise Ratio) statistics for all studied datasets are provided. We report the minimum, first quartile, average, median, third quartile, and maximum values. The best values are highlighted in bold.

4.3 Evaluation on a real painting

We now analyze the effectiveness of our method when employed in the context of a very common cultural heritage use case: the on-site free-form MLIC acquisition of large painted surfaces with subtle appearance and geometric details using a fixed camera and a hand-held moving light. The results analyzed in this article concern the outcomes of the capture, reconstruction, and relighting processes of two panels belonging to the retable of St. Bernardino (1455). This polyptych, originally housed in the chapel of St. Bernardino within the St. Francesco church in Cagliari, Italy, is currently held and exhibited at the Pinacoteca Nazionale in Cagliari. The first panel (refer to Fig. 23) measures 34×25cm and is executed in oil on a wooden support, portraying the prophet Daniel. The second panel (depicted in Fig. 26) is slightly larger, measuring 54×36 cm, and is adorned with a golden arched frame illustrating Christ in pity, upheld by an angel. The lower portion of the composition cannot be fully appreciated due to the state of preservation. Nevertheless, elements such as the Instruments of the Passion (Arma Christi) remain discernible, including the cross with the titulus INRI, the ladder for Christ's deposition, and the crucifixion nails. Both paintings were acquired in their pre-restoration state using a free-form setup, depicted in Fig. 22, that comprises a 36.3 Megapixel DSLR FX Nikon D810 Camera equipped with a 50 AF Nikkor Lens and a handheld white LED (5500K) covering the entire visible spectrum. Approximately 60 images were captured for each MLIC. The acquired data underwent calibration using four glossy spheres and a gray frame positioned around the object for estimating spatially-varying incident light direction and intensityemploying the camera and light calibration method introduced by Pintus et al. [41].

Here, our attention is directed towards three specific image subregions/crops of the first panel (refer to Fig. 23), chosen for their distinct characteristics. *Crop #0* encompasses a gold-colored relief featuring several cracks and

Dataset		Min	Q1	Avg	Median	Q3	Max
Paint-Texture-00	*Robust Trimmed PS*	0.961	0.977	0.981	0.980	0.987	0.993
	Ours	**0.978**	**0.988**	**0.989**	**0.990**	**0.991**	**0.994**
Paint-Texture-01	*Robust Trimmed PS*	0.956	0.986	0.987	0.989	0.992	0.996
	Ours	**0.964**	**0.991**	**0.991**	**0.994**	**0.994**	**0.997**
Paint-Texture-02	*Robust Trimmed PS*	0.962	0.986	0.987	0.989	0.991	0.994
	Ours	**0.975**	**0.990**	**0.992**	**0.994**	**0.995**	**0.996**
Mockup #0	*Robust Trimmed PS*	0.773	0.919	0.929	0.940	**0.957**	**0.972**
	Ours	**0.823**	**0.922**	**0.935**	**0.944**	0.955	0.970
Mockup #1	*Robust Trimmed PS*	0.756	0.919	0.923	0.939	0.954	**0.973**
	Ours	**0.828**	**0.926**	**0.934**	**0.942**	**0.956**	0.972
Mockup #2	*Robust Trimmed PS*	0.764	0.917	0.922	0.936	0.948	**0.967**
	Ours	**0.800**	**0.923**	**0.928**	**0.940**	**0.949**	**0.967**
Mockup #3	*Robust Trimmed PS*	0.785	0.966	0.960	0.980	**0.987**	**0.995**
	Ours	**0.829**	**0.970**	**0.966**	**0.982**	**0.987**	**0.995**
Mockup #4	*Robust Trimmed PS*	0.837	0.969	0.966	0.981	**0.991**	**0.996**
	Ours	**0.879**	**0.972**	**0.973**	**0.984**	**0.991**	**0.996**
Mockup #5	*Robust Trimmed PS*	0.832	0.970	0.966	**0.981**	**0.984**	**0.988**
	Ours	**0.871**	**0.972**	**0.970**	**0.981**	**0.984**	**0.988**
Retablo Crop #0	*Robust Trimmed PS*	0.700	0.859	0.911	0.952	0.977	0.988
	Ours	**0.825**	**0.891**	**0.936**	**0.958**	**0.980**	**0.993**
Retablo Crop #1	*Robust Trimmed PS*	0.313	0.566	0.794	0.923	0.981	0.991
	Ours	**0.564**	**0.709**	**0.860**	**0.937**	**0.986**	**0.993**
Retablo Crop #2	*Robust Trimmed PS*	0.445	0.714	0.860	0.979	0.996	**0.998**
	Ours	**0.734**	**0.846**	**0.928**	**0.987**	**0.997**	**0.998**

Table 3. *SSIM* (Structural Similarity Index) statistics for all studied datasets are provided. We report the minimum, first quartile, average, median, third quartile, and maximum values. The best values are highlighted in bold.

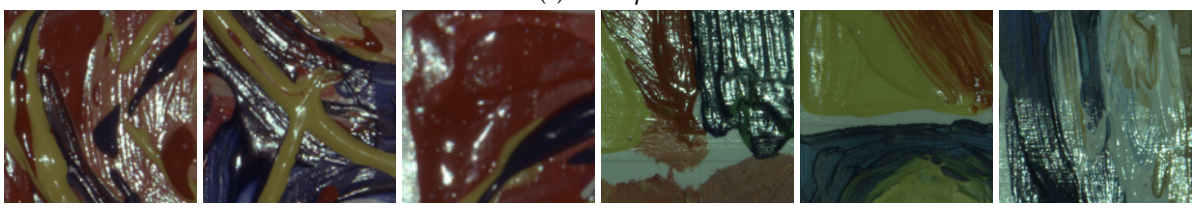
Synthetic Dataset	Method	ꟻLIP Statistics			
		Mean	Weighted Median	1st Quartile	3rd Quartile
Paint-Texture-00	*Robust Trimmed PS*	0.16	0.21	0.14	0.27
	Ours	**0.10**	**0.12**	**0.09**	**0.16**
Paint-Texture-01	*Robust Trimmed PS*	0.09	0.16	0.08	0.24
	Ours	**0.08**	**0.11**	**0.06**	**0.20**
Paint-Texture-02	*Robust Trimmed PS*	0.12	0.15	0.10	0.22
	Ours	**0.10**	**0.12**	**0.09**	**0.15**

Table 4. ꟻLIP statistics for the three synthetic datasets in Fig. 10.

a flat area with black writing. Typically, dark signals induce distortions in both the normal map and the final relighted image due to the non-negligible effect of acquisition noise. *Crop #1* depicts a decorative detail situated within a predominantly flat area with cracks, while *Crop #2* focuses on a facial detail within a highly glossy region. In most original images, the highlights are so pronounced that the painted face is obscured and overshadowed by the glossy signal, with evident paint mixing and various visible damages.

All evaluation metrics applied to these examples consistently demonstrate that shading normal computation significantly increases relighting quality (refer to Fig. 24). Some images exhibit a relative improvement in PSNR of over 20%, while SSIM performance shows even greater enhancements, reaching up to an 80% improvement. As observed in our other use cases, the improvement is particularly noticeable in front-illuminated images, where highlights or other glossy effects are prevalent.

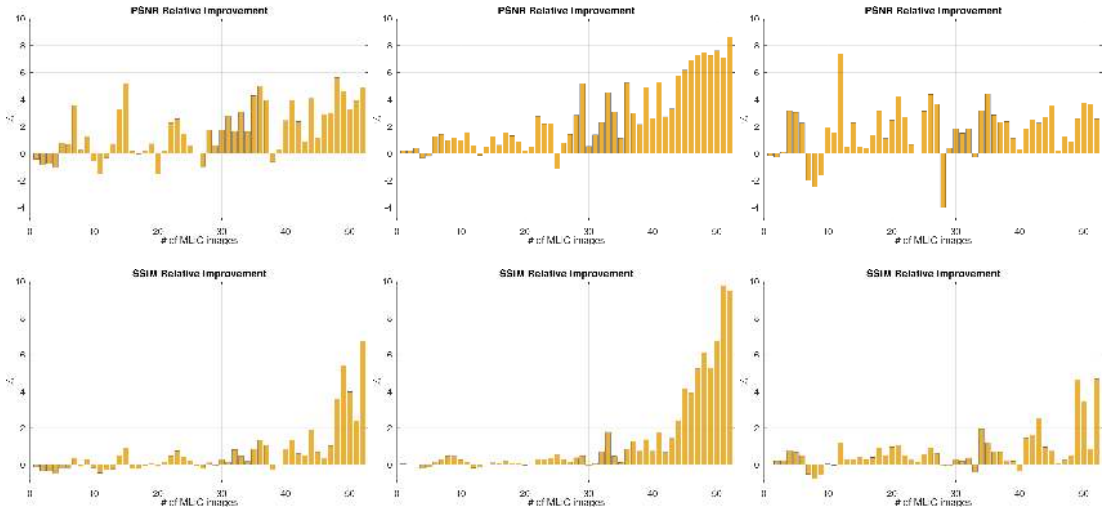
(a) *Mockups*



(b) *Mockup #0* (c) *Mockup #1* (d) *Mockup #2* (e) *Mockup #3* (f) *Mockup #4* (g) *Mockup #5*

Fig. 17. A set of painting mockups obtained through an RTI dome acquisition setup. Six image excerpts featuring diverse material and geometric properties serve as test cases to evaluate the effectiveness of the proposed normal refinement technique.

Fig. 25 depicts the rendered images alongside the ꟻLIP maps and associated statistics. Additionally, Tab. 6 summarizes in tabular form the numerical values of the mean, weighted median, 1st, and 3rd quartile of the error. In *Crop #0*, the improvement is evident as the gloss gold region is restored with greater accuracy, while the perceptual error in the flat region, particularly in the darker parts of the writing, is notably reduced. Similar results are observed in *Crop #1* and *Crop #2*, where both the error images and the plots illustrate how the normal refinement significantly enhances the quality of surface characterization and the corresponding virtual relighting. Across all tested image crops, the proposed normal refinement leads to an improved combination of surface geometry and material, resulting in relightable images with significantly reduced perceptual error. In this less controlled but more realistic setup, the enhancement provided by our method is notably more pronounced and evident than in synthetic tests and laboratory settings.



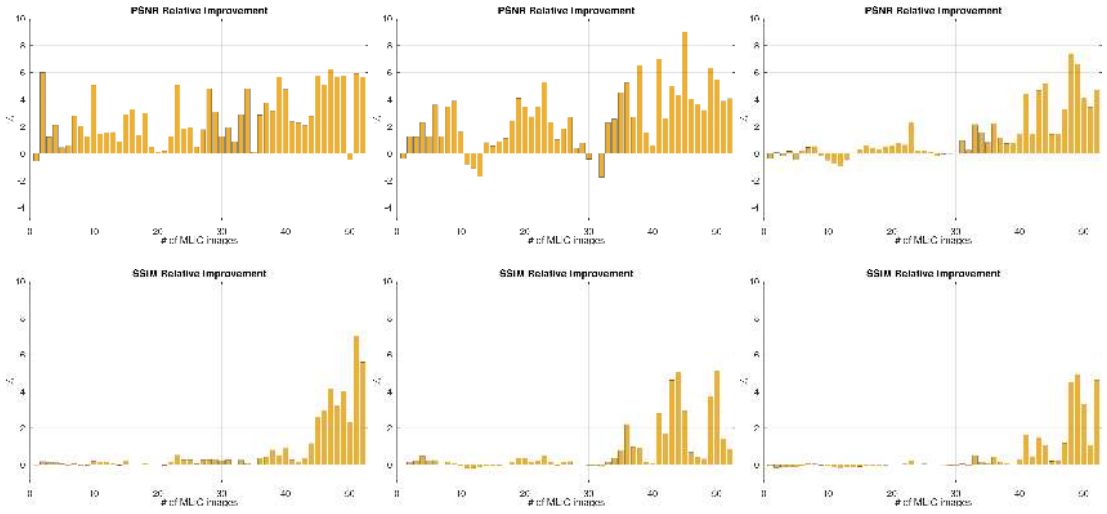
(a) *Mockup # 0* (b) *Mockup # 1* (c) *Mockup # 2*

Fig. 18. The performance of all relit images for the first set of three painting mockups is assessed. We present the percentage relative improvement for both the *PSNR* (Peak Signal-to-Noise Ratio) in the first row and the *SSIM* (Structural Similarity Index) in the second row. Noticeably better results are observed in images with frontal illumination (higher image numbers) compared to those with raking light conditions (lower image numbers), which primarily display the surface's diffuse signal.

Synthetic Dataset	Method	ꟻLIP Statistics			
		Mean	Weighted Median	1st Quartile	3rd Quartile
Mockup #0	*Robust Trimmed PS*	0.28	0.36	0.24	0.48
	Ours	**0.27**	**0.34**	**0.23**	**0.45**
Mockup #1	*Robust Trimmed PS*	0.24	0.32	0.20	0.46
	Ours	**0.20**	**0.26**	**0.17**	**0.37**
Mockup #2	*Robust Trimmed PS*	0.25	0.33	**0.20**	0.48
	Ours	**0.24**	**0.32**	**0.20**	**0.47**
Mockup #3	*Robust Trimmed PS*	0.30	0.38	0.28	0.51
	Ours	**0.27**	**0.35**	**0.25**	**0.47**
Mockup #4	*Robust Trimmed PS*	0.24	0.30	0.20	0.44
	Ours	**0.21**	**0.24**	**0.17**	**0.33**
Mockup #5	*Robust Trimmed PS*	0.20	0.33	0.16	0.58
	Ours	**0.18**	**0.28**	**0.14**	**0.48**

Table 5. ꟻLIP statistics for the six real-world mockups in Fig. 17.

As a last additional consideration, another qualitative aspect that aids in assessing the possibly higher accuracy of the extracted normal map is the degree of visual separation between the geometric signal and the material, color, or optical maps. The tendency for some information to leak into others is a recognized challenge, influenced by various factors including the inherent imperfections of analytical models of image formation, as well as non-ideal and challenging-to-measure effects such as self and cast shadows, interreflections, and global lighting effects. Fig. 26 and Fig. 27 showcase examples where it is evident that the refined normal maps exhibit fewer artifacts (and thus less distortion) compared to theoriginal normal map computed with the *Robust Trimmed PS*. Notably, the color signal of the wood knot in the cross remains visible in the non-refined normal map. Our



(a) *Mockup # 3* (b) *Mockup # 4* (c) *Mockup # 5*

Fig. 19. The performance of all relit images for the second set of three painting mockups is assessed. We present the percentage relative improvement for both the *PSNR* (Peak Signal-to-Noise Ratio) in the first row and the *SSIM* (Structural Similarity Index) in the second row. Noticeably better results are observed in images with frontal illumination (higher image numbers) compared to those with raking light conditions (lower image numbers), which primarily display the surface's diffuse signal.

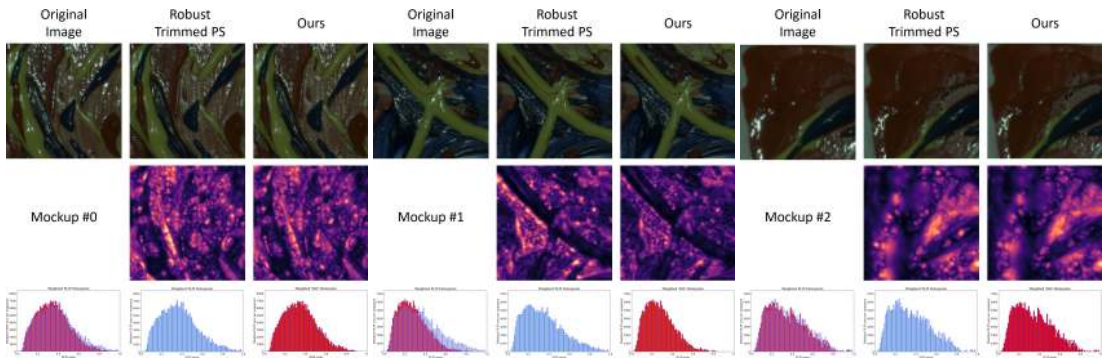


Fig. 20. We juxtapose selected original images of the first set of three painting mockups with their digitally relit versions generated from the SV-BRDF calculated using a normal map from a standard *Robust Trimmed PS* and our refined normal map. The final two rows showcase the respective ꟻLIP maps and associated statistics. The leftmost plot in the third row of each group of images overlays the two sets of statistics to highlight the improvement more effectively.

method effectively homogenizes this region, causing the color detail to disappear in the geometric reconstruction. The rightmost images in the figure display the angular difference, highlighting the areas where the correction has been more pronounced. As we can see, several corrections appear in areas with colored details. Similarly, in the other Retablo (refer to Fig. 27), black drawings and golden sections appear to significantly distort the normal

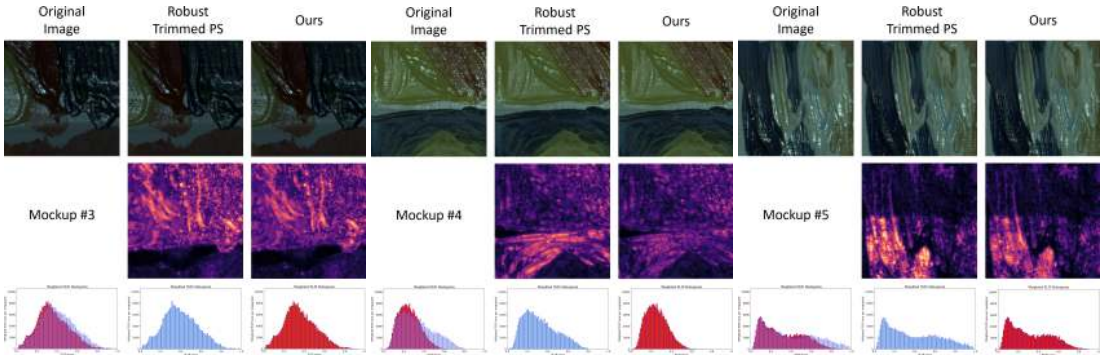


Fig. 21. We juxtapose selected original images of the second set of three painting mockups with their digitally relit versions generated from the SV-BRDF calculated using a normal map from a standard *Robust Trimmed PS* and our refined normal map. The final two rows showcase the respective ƎLIP maps and associated statistics. The leftmost plot in the third row of each group of images overlays the two sets of statistics to highlight the improvement more effectively.

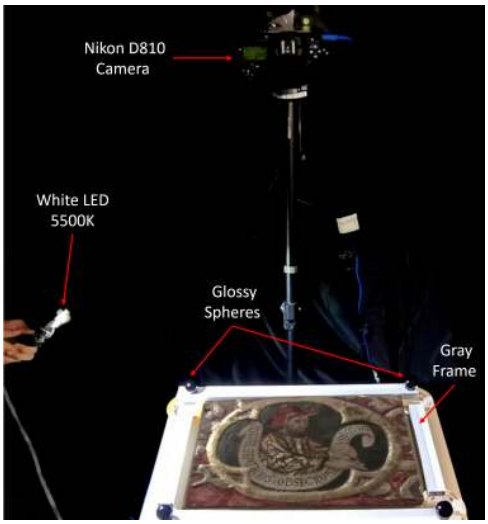


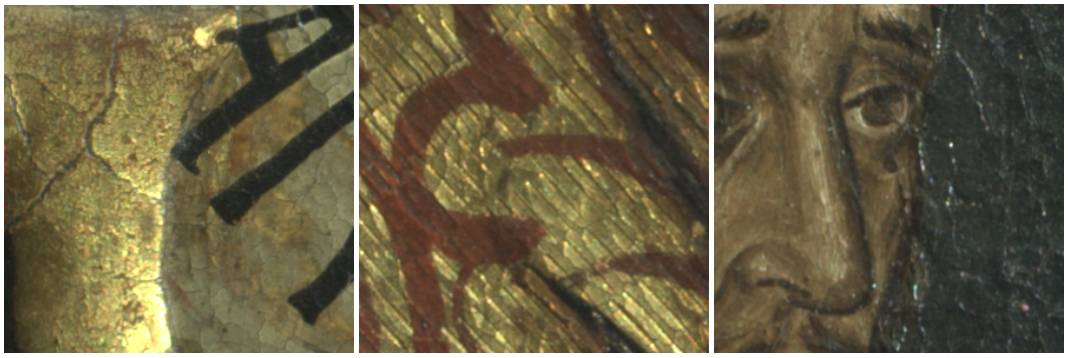
Fig. 22. Photographing and illuminating a painting (panel of the polyptych retable of Saint Bernardino, 1455, Cagliari, Italy). The painting's optical properties under varying illumination are captured by taking several dozen photos using a stationary reflex camera and a handheld LED.

map. Our normal refinement approach successfully eliminates the pattern caused by the drawings on the saint's clothes and compensates for deformations caused by the bright, golden regions surrounding the red motif.

5 Conclusions

Multi-Light Image Collections (MLICs) offer comprehensive data for characterizing imaged objects and are extensively used in many domains, among which cultural heritage is prominent. In particular, a variety of Photometric Stereo techniques can provide a fine geometric characterization of the imaged object's surface, offering a basis for many shape analysis applications [42]. For visual exploration, such normal maps are often

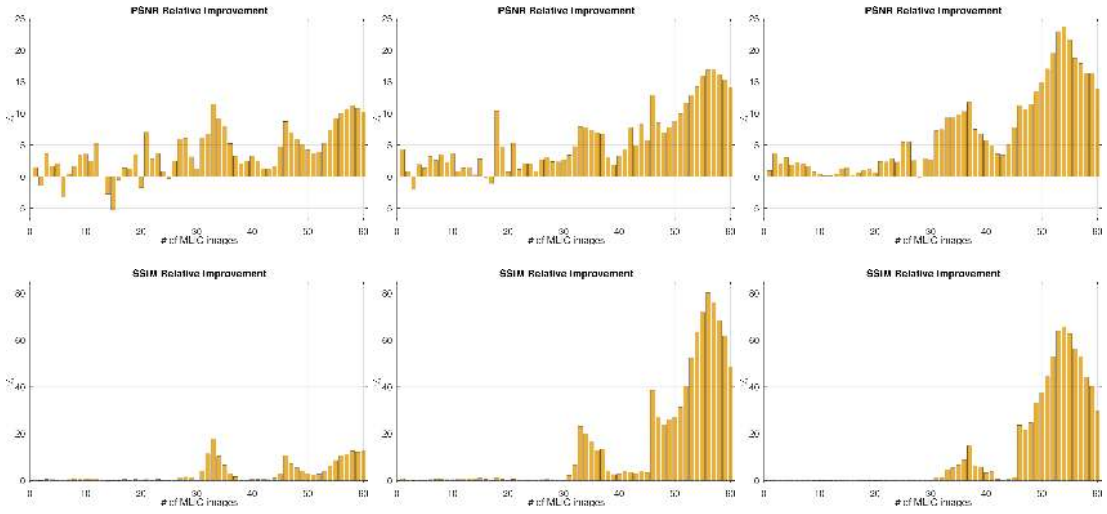
(a) *A painting (component of the Retable of S. Berndardino)*



(b) *Retablo Crop #0* (c) *Retablo Crop #1* (d) *Retablo Crop #2*

Fig. 23. An artwork (a panel from the Retable of Saint Bernardino) captured using a free-form RTI acquisition setup. Three image excerpts showcasing various material and geometric properties are employed to evaluate the effectiveness of the proposed normal refinement technique.

coupled with SV-BRDF maps, which provide per-pixel coefficients of analytical BRDF models, to create compact



(a) *Retablo Crop #0* (b) *Retablo Crop #1* (c) *Retablo Crop #2*

Fig. 24. The performance of all relit images for the three Retablo crops is assessed. We present the percentage relative improvement for both the *PSNR* (Peak Signal-to-Noise Ratio) in the first row and the *SSIM* (Structural Similarity Index) in the second row. Noticeably better results are observed in images with frontal illumination (higher image numbers) compared to those with raking light conditions (lower image numbers), which primarily display the surface's diffuse signal.

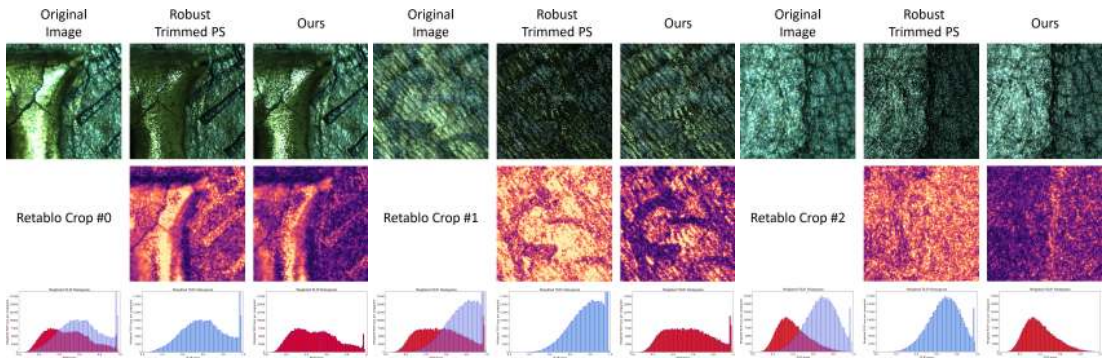


Fig. 25. We juxtapose selected original images of the three Retablo's crops with their digitally relit versions generated from the SV-BRDF calculated using a normal map from a standard *Robust Trimmed PS* and our refined normal map. The final two rows showcase the respective ꟻLIP maps and associated statistics. The leftmost plot in the third row of each group of images overlays the two sets of statistics to highlight the improvement more effectively.

and fast to evaluate models that can support a compelling visual exploration inside interactive relighting viewers. However, such an approach may often lead to suboptimal visual outcomes. This occurs not only when the geometric normals are inaccurately recovered by the Photometric Stereo method, but also in the presence of a "perfect" geometric reconstruction, due to a variety of practical reasons. These include the presence of non-local lighting effects in the input not taken into account in the BRDF fitter and, more in general, the discrepancy

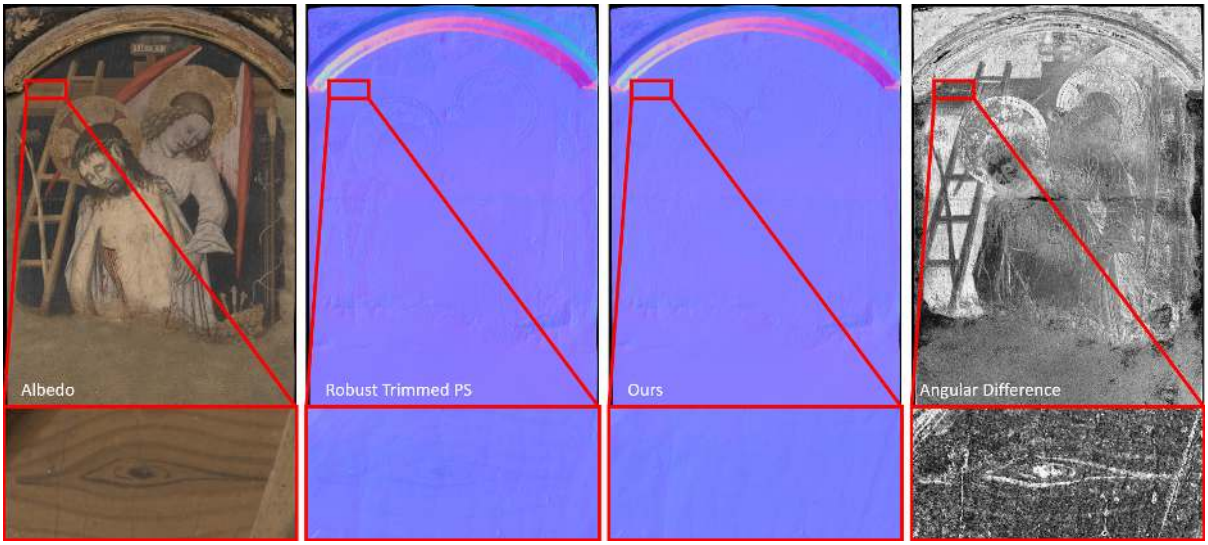


Fig. 26. Albedo to normal leakage is compensated by our normal refinement technique. Refined normal maps contain less spurious signal (and so less normal distortion) than the normal map computed with the *Robust Trimmed PS*. The rightmost images display the angular difference, emphasizing the areas where the corrections have been more pronounced.

Synthetic Dataset	Method	ꟻLIP Statistics			
		Mean	Weighted Median	1st Quartile	3rd Quartile
Retablo Crop #0	*Robust Trimmed PS*	0.53	0.61	0.46	0.76
	Ours	**0.39**	**0.47**	**0.32**	**0.63**
Retablo Crop #1	*Robust Trimmed PS*	0.75	0.82	0.68	0.94
	Ours	**0.45**	**0.54**	**0.37**	**0.71**
Retablo Crop #2	*Robust Trimmed PS*	0.64	0.69	0.58	0.79
	Ours	**0.35**	**0.38**	**0.29**	**0.50**

Table 6. ꟻLIP statistics for the three retablo crops in Fig. 23.

between the behavior of a pure normal+BRDF model and complex optical behaviors of many objects (e.g., due to layering, subsurface scattering, and other effects).

In this work, we have proposed to decouple the normals used for shading from the geometric normals, as done in many rendering frameworks. Building on the common and general assumption that the reflectance under local illumination in a typical normal+BRDF model is a monotonic function of the angle between the shading normal and the bisector of lighting and viewing directions, we have proposed a solution that corrects the orientation of the geometric normals into shading normals that better meet this assumption. Differently from solutions that force the usage of a coupled shape and material solver, our method integrates seamlessly into existing reconstruction pipelines, supporting a free mixing and matching of Photometric Stereo methods, BRDF models, and BRDF fitters. Users can decide, in their application, to use our corrected normals also as improved geometric normals, especially when using the most common normal recovery techniques (e.g., based on trimming). As a more common alternative, they may opt for using separate normal maps for geometric analysis and visualization. Through synthetic, laboratory, and real-world painting data, we have shown our method's effectiveness in enhancing the recovery of the original illumination at different angles in relighting applications.

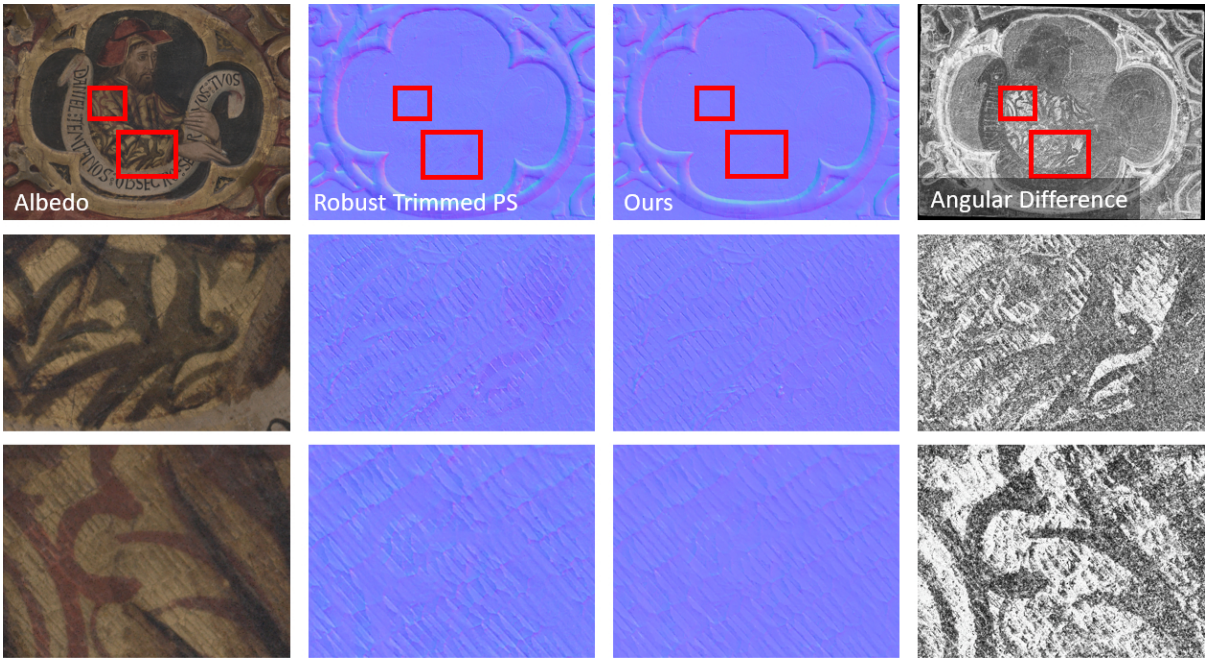


Fig. 27. Albedo to normal leakage is compensated by our normal refinement technique. Refined normal maps contain less spurious signal (and so less normal distortion) than the normal map computed with the *Robust Trimmed PS*. The rightmost images display the angular difference, emphasizing the areas where the corrections have been more pronounced.

Our current work is concentrating on further enhancing the approach, in particular by reducing its computation time. While latency in producing results is not critical, as the method is executed at processing time and often coupled with other demanding steps for normal and material recovery, we still think that there is a large margin of improvement in its execution time. In particular, we plan to improve the evaluation cost of the individual metric evaluation steps, avoiding the sorting of the appearance profile at each evaluation. We are also looking at employing an adaptive global solver, rather than the fixed-cost coarse-to-fine approach discussed in this paper. Finally, we plan to reduce the number of fully computed pixels by sharing solutions between nearby areas.

Due to its simplicity and modularity, we expect this method to be adopted to improve the quality of the existing reconstruction pipelines.

Moreover, our current work has been, so far, tested in the context of the main target of MLIC acquisition and reconstruction in cultural heritage, which are objects with mostly flat global scale geometry but intricate small-scale details. This class of objects includes not only paintings, such as those that we have evaluated in our results section but also other cultural objects such as frescoes, graffiti, wall incisions, coins, bas-reliefs, and many more. The technique itself is, however, not limited to this use case, and we plan to evaluate its suitability for 3D objects with complex geometries, exhibiting a variety of curvatures and occlusions from different angles.

Acknowledgments

We thank the National Archaeological Museum and the National Art Gallery in Cagliari for access to artworks for digitization and fruitful collaboration. This work received funding from Sardinian regional authorities under the XDATA project (art 9 L.R. 20/2015). We also acknowledge the contribution of project “REFLEX – REFLectance

EXploration: improving the acquisition, distribution, and exploration of multi-light image collections for surface characterization and analysis" (PRIN2022) funded by the EU Next-GenerationEU PNRR, component M4C2-1.1.

References

[1] Jens Ackermann, Michael Goesele, et al. 2015. A Survey of Photometric Stereo Techniques. *Foundations and Trends in Computer Graphics and Vision* 9, 3-4 (2015), 149–254. doi:10.1561/0600000065

[2] Miika Aittala, Tim Weyrich, and Jaakko Lehtinen. 2013. Practical SVBRDF Capture in the Frequency Domain. *ACM TOG* 32, 4 (2013), 110–1. doi:10.1145/2461912.2461978

[3] Neil Alldrin, Todd Zickler, and David Kriegman. 2008. Photometric Stereo with Non-Parametric and Spatially-Varying Reflectance. In *Proc. CVPR*. IEEE, 1–8. doi:10.1109/CVPR.2008.4587656

[4] Pontus Andersson, Jim Nilsson, Tomas Akenine-Möller, Magnus Oskarsson, Kalle Åström, and Mark D. Fairchild. 2020. ꟻLIP: A Difference Evaluator for Alternating Images. *Proc. ACM SIGGRAPH* 3, 2 (2020), 15:1–15:23. doi:10.1145/3406183

[5] Jonathan T Barron and Jitendra Malik. 2014. Shape, Illumination, and Reflectance from Shading. *IEEE TPAMI* 37, 8 (2014), 1670–1687. doi:10.1109/TPAMI.2014.2377712

[6] Ronen Basri, David Jacobs, and Ira Kemelmacher. 2007. Photometric Stereo with General, Unknown Lighting. *International Journal of Computer Vision* 72, 3 (2007), 239–257. doi:10.1007/s11263-006-8815-7

[7] Guanying Chen, Kai Han, Boxin Shi, Yasuyuki Matsushita, and Kwan-Yee K Wong. 2019. Self-Calibrating Deep Photometric Stereo Networks. In *Proc. CVPR*. IEEE, 8739–8747. doi:10.48550/arXiv.1903.07366

[8] Zhe Chen, Shohei Nobuhara, and Ko Nishino. 2021. Invertible Neural BRDF for Object Inverse Rendering. *IEEE TPAMI* 44, 12 (2021), 9380–9395. doi:10.1109/TPAMI.2021.3129537

[9] Gyeongmin Choe, Srinivasa G Narasimhan, and In So Kweon. 2016. Simultaneous Estimation of Near IR BRDF and Fine-Scale Surface Geometry. In *Proc. CVPR*. IEEE, 2452–2460. doi:10.1109/CVPR.2016.269

[10] Julie Dorsey, Holly Rushmeier, and François Sillion. 2010. *Digital Modeling of Material Appearance*. Elsevier.

[11] Mark S Drew, Yacov Hel-Or, Tom Malzbender, and Nasim Hajari. 2012. Robust Estimation of Surface Properties and Interpolation of Shadow/Specularity Components. *Image and Vision Computing* 30, 4-5 (2012), 317–331. doi:10.1016/j.imavis.2012.02.012

[12] Shireen Y Elhabian, Ham Rara, and Aly A Farag. 2011. Towards Accurate and Efficient Representation of Image Irradiance of Convex-Lambertian Objects under Unknown Near Lighting. In *Proc. ICCV*. IEEE, 1732–1737. doi:10.1109/ICCV.2011.6126437

[13] Every Texture. 2024. 3D Texture Database. https://everytexture.com/ [Online; accessed-July-31-2024].

[14] Duan Gao, Xiao Li, Yue Dong, Pieter Peers, Kun Xu, and Xin Tong. 2019. Deep Inverse Rendering for High-Resolution SVBRDF Estimation from an Arbitrary Number of Images. *ACM TOG* 38, 4 (2019), 134–1. doi:10.1145/3306346.3323042

[15] Athinodoros Georghiades. 2003. Incorporating the Torrance and Sparrow Model of Reflectance in Uncalibrated Photometric Stereo. In *Proc. ICCV*. IEEE, 816–823. doi:10.1109/ICCV.2003.1238432

[16] D.B. Goldman, B. Curless, A. Hertzmann, and S.M. Seitz. 2005. Shape and Spatially-Varying BRDFs from Photometric Stereo. In *Proc. ICCV*. IEEE, 341–348 Vol. 1. doi:10.1109/ICCV.2005.219

[17] Geoffrey Grimmett. 2006. *The Random-Cluster Model*. Vol. 333. Springer. doi:10.1007/978-3-540-32891-9

[18] Dar'ya Guarnera, Giuseppe Claudio Guarnera, Abhijeet Ghosh, Cornelia Denk, and Mashhuda Glencross. 2016. BRDF Representation and Acquisition. *Computer Graphics Forum* 35, 2 (2016), 625–650. doi:10.1111/cgf.12867

[19] Aaron Hertzmann and Steven M Seitz. 2005. Example-based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs. *IEEE TPAMI* 27, 8 (2005), 1254–1264. doi:10.1109/TPAMI.2005.158

[20] Tomoaki Higo, Yasuyuki Matsushita, and Katsushi Ikeuchi. 2010. Consensus Photometric Stereo. In *Proc. CVPR*. IEEE, 1157–1164. doi:10.1109/CVPR.2010.5540084

[21] Michael Holroyd, Jason Lawrence, Greg Humphreys, and Todd Zickler. 2008. A Photometric Approach for Estimating Normals and Tangents. *ACM TOG* 27, 5 (2008), 1–9. doi:10.1145/1457515.1409086

[22] Xinyi Huang, Hao Cao, and Bingjing Jia. 2023. Optimization of Levenberg Marquardt Algorithm Applied to Nonlinear Systems. *Processes* 11, 6 (2023), 1794. doi:10.3390/pr11061794

[23] Zhuo Hui and Aswin C Sankaranarayanan. 2015. A Dictionary-Based Approach for Estimating Shape and Spatially-Varying Reflectance. In *Proc. ICCP*. IEEE, 1–9. doi:10.48550/arXiv.1503.04265

[24] Zhuo Hui and Aswin C Sankaranarayanan. 2016. Shape and Spatially-Varying Reflectance Estimation from Virtual Exemplars. *IEEE TPAMI* 39, 10 (2016), 2060–2073. doi:10.1109/TPAMI.2016.2623613

[25] Satoshi Ikehata. 2018. CNN-PS: CNN-based Photometric Stereo for General Non-Convex Surfaces. In *Proc. ECCV*. Springer, 3–18. doi:10.1007/978-3-030-01267-0_1

[26] Satoshi Ikehata, David Wipf, Yasuyuki Matsushita, and Kiyoharu Aizawa. 2012. Robust Photometric Stereo using Sparse Regression. In *Proc. CVPR*. IEEE, 318–325. doi:10.1109/CVPR.2012.6247691

[27] Yakun Ju, Kin-Man Lam, Wuyuan Xie, Huiyu Zhou, Junyu Dong, and Boxin Shi. 2024. Deep Learning Methods for Calibrated Photometric Stereo and Beyond: A Survey. *arXiv preprint arXiv:2212.08414* (2024), 1–19. doi:10.48550/arXiv.2212.08414

[28] Guillaume Lavoué, Nicolas Bonneel, Jean-Philippe Farrugia, and Cyril Soler. 2021. Perceptual Quality of BRDF Approximations: Dataset and Metrics. *Computer Graphics Forum* 40, 2 (2021), 327–338. doi:10.1111/cgf.142636

[29] Jason Lawrence, Aner Ben-Artzi, Christopher DeCoro, Wojciech Matusik, Hanspeter Pfister, Ravi Ramamoorthi, and Szymon Rusinkiewicz. 2006. Inverse Shade Trees for Non-Parametric Material Representation and Editing. *ACM TOG* 25, 3 (2006), 735–745. doi:10.1145/1141911.1141949

[30] Chen Liu, Michael Fischer, and Tobias Ritschel. 2023. Learning to Learn and Sample BRDFs. *Computer Graphics Forum* 42, 2 (2023), 201–211. doi:10.1111/cgf.14754

[31] M.I.A. Lourakis. Jul. 2004. levmar: Levenberg-Marquardt nonlinear least squares algorithms in C/C++. [web page] http://users.ics.forth.gr/ lourakis/levmar/. [accessed-July-31-2024].

[32] Tom Malzbender, Dan Gelb, and Hans Wolters. 2001. Polynomial Texture Maps. *Proc. ACM SIGGRAPH* (2001), 519–528. doi:10.1145/383259.383320

[33] MathWorks. 2024. monotonicity, Quantify monotonic trend in condition indicators. https://www.mathworks.com/help/predmaint/ref/monotonicity.html [Online; accessed-July-31-2024].

[34] Daisuke Miyazaki, Kenji Hara, and Katsushi Ikeuchi. 2010. Median Photometric Stereo as Applied to the Segonko Tumulus and Museum Objects. *International Journal of Computer Vision* 86 (2010), 229–242. doi:10.1007/s11263-009-0262-9

[35] Yasuhiro Mukaigawa, Yasunori Ishii, and Takeshi Shakunaga. 2007. Analysis of Photometric Factors Based on Photometric Linearization. *JOSA* 24, 10 (2007), 3326–3334. doi:10.1364/JOSAA.24.003326

[36] László Neumann, Attila Neumannn, and László Szirmay-Kalos. 1999. Compact Metallic Reflectance Models. *Computer Graphics Forum* 18, 3 (1999), 161–172. doi:10.1111/1467-8659.00337

[37] Taishi Ono, Hiroyuki Kubo, Kenichiro Tanaka, Takuya Funatomi, and Yasuhiro Mukaigawa. 2019. Practical BRDF Reconstruction Using Reliable Geometric Regions from Multi-View Stereo. *Computational Visual Media* 5 (2019), 325–336. doi:10.1007/s41095-019-0150-3

[38] Ruggero Pintus, Moonisa Ahsan, Antonio Zorcolo, Fabio Bettio, Fabio Marton, and Enrico Gobbetti. 2023. Exploiting Local Shape and Material Similarity for Effective SV-BRDF Reconstruction from Sparse Multi-Light Image Collections. *ACM JOCCH* 16, 2 (2023), 39:1–39:31. doi:10.1145/3593428

[39] Ruggero Pintus, Tinsae Gebrechristos Dulecha, Irina Ciortan, Enrico Gobbetti, and Andrea Giachetti. 2019. State-of-the-art in Multi-Light Image Collections for Surface Visualization and Analysis. *Computer Graphics Forum* 38, 3 (2019), 909–934. doi:10.1111/cgf.13732

[40] Ruggero Pintus, Andrea Giachetti, Giovanni Pintore, and Enrico Gobbetti. 2017. Guided Robust Matte-Model Fitting for Accelerating Multi-light Reflectance Processing Techniques. In *Proc. British Machine Vision Conference*. BMVA, 32.1–32.15. doi:10.5244/C.31.32

[41] Ruggero Pintus, Alberto Jaspe Villanueva, Antonio Zorcolo, Markus Hadwiger, and Enrico Gobbetti. 2021. A Practical and Efficient Model for Intensity Calibration of Multi-Light Image Collections. *The Visual Computer* 37, 9 (September 2021), 2755–2767. doi:10.1007/s00371-021-02172-9

[42] Ruggero Pintus, Kazim Pal, Ying Yang, Tim Weyrich, Enrico Gobbetti, and Holly Rushmeier. 2016. A Survey of Geometric Analysis in Cultural Heritage. *Computer Graphics Forum* 35, 1 (2016), 4–31. doi:10.1111/cgf.12668

[43] Gilles Pitard, Gaëtan Le Goïc, Alamin Mansouri, Hugues Favrelière, Simon-Frederic Desage, Serge Samper, and Maurice Pillet. 2017. Discrete Modal Decomposition: a New Approach for the Reflectance Modeling and Rendering of Real Surfaces. *Machine Vision and Applications* 28, 5-6 (2017), 607–621. doi:10.1007/s00138-017-0856-0

[44] Peiran Ren, Yue Dong, Stephen Lin, Xin Tong, and Baining Guo. 2015. Image Based Relighting using Neural Networks. *ACM TOG* 34, 4 (2015), 111:1–111:12. doi:10.1145/2766899

[45] Leonardo Righetto, Fabio Bettio, Federico Ponchio, Andrea Giachetti, and Enrico Gobbetti. 2023. Effective Interactive Visualization of Neural Relightable Images in a Web-based Multi-Layered Framework. In *Proc. GCH*. Eurographics, 57–66. doi:10.2312/gch.20231158

[46] Leonardo Righetto, Arianna Traviglia, Michela De Bernardin, Enrico Gobbetti, Federico Ponchio, and Andrea Giachetti. 2023. Ancient Coins' Surface Inspection with Web-Based Neural RTI Visualization. In *Optics for Arts, Architecture, and Archaeology (O3A) IX*, Vol. 12620. SPIE, 91–98. doi:10.1117/12.2674888

[47] Fabiano Romeiro and Todd Zickler. 2010. Blind Reflectometry. In *Proc. ECCV*. Springer, 45–58. doi:10.1007/978-3-642-15549-9_4

[48] Roland Ruiters and Reinhard Klein. 2009. Heightfield and Spatially Varying BRDF Reconstruction for Materials with Interreflections. *Computer Graphics Forum* 28, 2 (2009), 513–522. doi:10.1111/j.1467-8659.2009.01390.x

[49] Szymon M Rusinkiewicz. 1998. A New Change of Variables for Efficient BRDF Representation. In *Rendering Techniques*. Springer, 11–22. doi:10.1007/978-3-7091-6453-2_2

[50] Hiroaki Santo, Masaki Samejima, Yusuke Sugano, Boxin Shi, and Yasuyuki Matsushita. 2017. Deep Photometric Stereo Network. In *Proc. ICCV workshops*. IEEE, 501–509. doi:10.1109/ICCVW.2017.66

[51] Boxin Shi, Ping Tan, Yasuyuki Matsushita, and Katsushi Ikeuchi. 2012. Elevation Angle from Reflectance Monotonicity: Photometric Stereo for General Isotropic Reflectances. In *Proc. ECCV*. Springer, 455–468. doi:10.1007/978-3-642-33712-3_33

[52] Boxin Shi, Ping Tan, Yasuyuki Matsushita, and Katsushi Ikeuchi. 2013. Bi-Polynomial Modeling of Low-Frequency Reflectances. *IEEE TPAMI* 36, 6 (2013), 1078–1091. doi:10.1109/TPAMI.2013.196

[53] Boxin Shi, Zhe Wu, Zhipeng Mo, Dinglong Duan, Sai-Kit Yeung, and Ping Tan. 2016. A Benchmark Dataset and Evaluation for Non-Lambertian and Uncalibrated Photometric Stereo. In *IEEE TPAMI*. IEEE, 3707–3716. doi:10.1109/TPAMI.2018.2799222

[54] William M Silver. 1980. *Determining Shape and Reflectance using Multiple Images*. Ph. D. Dissertation. Massachusetts Institute of Technology.

[55] David Tingdahl, Christoph Godau, and Luc Van Gool. 2012. Base Materials for Photometric Stereo. In *Proc. ECCV*. Springer, 350–359. doi:10.1007/978-3-642-33868-7_35

[56] Tanaboon Tongbuasirilai, Jonas Unger, Christine Guillemot, and Ehsan Miandji. 2022. A Sparse Non-Parametric BRDF Model. *ACM TOG* 41, 5 (2022), 1–18. doi:10.1145/3533427

[57] Kenneth E Torrance and Ephraim M Sparrow. 1967. Theory for Off-Specular Reflection from Roughened Surfaces. *JOSA* 57, 9 (1967), 1105–1114. doi:10.1364/JOSA.57.001105

[58] Bruce Walter. 2005. *Notes on the Ward BRDF*. Technical Report PCG-05. Program of Computer Graphics, Cornell University.

[59] Feishi Wang, Jieji Ren, Heng Guo, Mingjun Ren, and Boxin Shi. 2023. DiLiGenT-Pi: Photometric Stereo for Planar Surfaces with Rich Details-Benchmark Dataset and Beyond. In *Proc. ICCV*. IEEE, 9477–9487. doi:10.1109/ICCV51070.2023.00869

[60] Robert J Woodham. 1980. Photometric Method for Determining Surface Orientation from Multiple Images. *Optical engineering* 19, 1 (1980), 139–144. doi:10.1117/12.7972479

[61] Lun Wu, Arvind Ganesh, Boxin Shi, Yasuyuki Matsushita, Yongtian Wang, and Yi Ma. 2011. Robust Photometric Stereo via Low-Rank Matrix Completion and Recovery. In *Proc. ACCV*. Springer, 703–717. doi:10.1007/978-3-642-19318-7_55

[62] Tai-Pang Wu and Chi-Keung Tang. 2009. Photometric Stereo via Expectation Maximization. *IEEE TPAMI* 32, 3 (2009), 546–560. doi:10.1109/TPAMI.2009.15

[63] X-Rite. 2020. Appearance Exchange Format. https://www.xrite.com/axf [Online; accessed-July-31-2024].

[64] Zexiang Xu, Kalyan Sunkavalli, Sunil Hadap, and Ravi Ramamoorthi. 2018. Deep Image-based Relighting from Optimal Sparse Samples. *ACM TOG* 37, 4 (2018), 126. doi:10.1145/3197517.3201313

[65] Mingjing Zhang and Mark S Drew. 2014. Efficient Robust Image Interpolation and Surface Properties using Polynomial Texture Mapping. *EURASIP Journal on Image and Video Processing* 2014, 1 (2014), 25. doi:10.1186/1687-5281-2014-25

[66] Todd Zickler, Ravi Ramamoorthi, Sebastian Enrique, and Peter N. Belhumeur. 2006. Reflectance Sharing: Predicting Appearance from a Sparse Set of Images of a Known Shape. *IEEE TPAMI* 28, 8 (2006), 1287–1302. doi:10.1109/TPAMI.2006.170