DATA-DRIVEN HOMOGENEOUS APPEARANCE MODELING FROM A SINGLE IMAGE UNDER UNCONTROLLED NATURAL LIGHTING

Victoria Cooper

James Bieron

Pieter Peers

The College of William & Mary

March 24, 2019

0   Abstract

We present a novel method for modeling the reflectance of a convex object under known natural lighting from a single image. We utilize the knowledge that similar materials often have similar reflectance functions. This allows us to model reflectance functions using Gaussian mixture models where each Gaussian is a material type. The solution is found using maximum a-posteriori (MAP) and non-linear optimizations to select the best option from our set of solutions. We demonstrate the effectiveness of this approach using a database of measured reflectance functions and under multiple natural lighting environments.

1   Introduction

When modeling reflectance, reflectance functions such as Bidirectional Reflectance Distribution Functions (BRDF) are used to express the reflectance of a material. Data-driven appearance models[3,4] expresses the (BRDF) of a homogeneous material as a linear combination of a large set of measured *"basis"* BRDFs. The key assumption is that this large set of basis BRDFs covers the full space of BRDFs, and any BRDF in this space, can be represented as convex combination of these basis BRDFs, thereby inheriting all the intricate reflectance details present in the measured basis BRDFs. Recent advances have shown great promise in reconstructing a data-driven BRDF from very few measurements.[8,11] However, these methods rely on controlled directional or point lighting, and do not generalize to uncontrolled natural lighting. A key problem in generalizing prior methods to natural lighting is that these prior methods rely on a non-linear encoding (e.g., logarithmic) to compress the dynamic range of the the basis BRDFs in order to regularize the estimation of the model parameters. Such non-linear encoding can only be undone after parameter estimation if the observations consist of direct BRDF observations (i.e., a single view and a single light direction per observation). In contrast, observations under natural lighting are the result of an integration of the BRDF times lighting over all directions, and only linear transformations of the BRDF are transparent to this integration.

In this paper we present a method for reconstructing the BRDF of a homogeneous material from a single photograph of a convex object

with known shape and observed under a known, but uncontrolled, natural lighting condition. Our method does not rely on a non-linear transformation of the basis BRDFs. Instead we directly operate on the unmodified BRDFs. To regularize the estimation of the model parameters we leverage the reflectance similarities between BRDFs in a material class. We approximate the space of homogeneous BRDFs with a Gaussian mixture model. Each normal distribution in the Gaussian mixture model represents a material class, and we assign each basis material to the class with the highest likelihood. We formulate the estimate of the model parameters as a maximum a-posteriori optimization that maximizes the likelihood that the model parameters explain the observations, as well as the likelihood that the model belongs to the material class. We exploit the observation that in high dimensional spaces everything is distant, and approximate the maximum a-posteriori optimization by a efficient linear least squares approximation per material class. Finally, we select the most likely provisional least squares solution based on the maximum a-posteriori error. We demonstrate the efficacy of our solution using the MERL BRDF database under a variety of natural lighting conditions.

Recovering realistic reflectance of materials has several direct applications to NASA. Both the Human Exploration and Outreach directive[6] and the Science mission directorate,[7] show NASA's desire to explore and learn more about our solar system. This includes colonizing Mars and exploring the planets in our solar system. Sending astronauts on missions into space is dangerous, so to prepare astronauts need to be well trained for all potential emergencies that may occur on their mission. Virtual reality, VR, is a tool that creates a realistic and safe environment for the astronauts to train in. So for this training to be useful, they must look and feel realistic to the astronauts. A study performed by Seymour et al.[10] found that sur-

geons trained using VR were more efficient and accurate than their counterparts. Additionally, another goal of NASA's is to educate the public through outreach programs.[5] Recovering realistic reflectance properties can be used to enhance learning experiences through VR simulations. This can be used by researches to digitize artifacts that can then be studied. Alternatively, these digital artifacts can be put on display to educate the public on NASA's goals and missions.

2   Overview

2.1   Data-driven BRDF

The reflectance behavior of a homogeneous material is described by the bidirectional reflectance distribution function (BRDF) $\rho(\omega_i, \omega_o)$: a $4D$ function defined as the ratio of incident irradiance for an incident direction ω_i over the outgoing radiance for outgoing directions ω_o.

In this paper, we follow the data-driven BRDF model of Matusik et al.[4] that characterizes the BRDF ρ as a linear combination of a large set of n measured materials $b_i, i \in [1, n]$. The underlying idea is that the set of measured BRDFs spans the space of BRDFs, and any material's BRDF should lie in this space:

$$\rho = Bw, \tag{1}$$

where we stack the BRDF ρ and basis BRDFs b_i in a vector of length p, and form the matrix B by stacking each basis vector in a column: $B = [b_1, ..., b_n]$. The model parameters are stacked in a vector w of n scalar weights. We directly use the BRDF parameterization of the MERL BRDF database,[3] and $p = 90 \times 90 \times 180$. Furthermore, similar as in Nielsen et al.,[8] we consider each color channel of the 100 MERL BRDFs as a basis BRDF, and thus $n = 300$.

Due to the large dynamic range between specular peaks versus diffuse reflectance, prior

work[3,8,11] has applied a non-linear compression function ζ to make the estimation of w less sensitive to errors on the (large) specular peaks:

$$\rho' = B'w', \qquad (2)$$

where $B' = [\zeta(b_1), ..., \zeta(b_n)]$. An expansion ζ^{-1} is applied to the compressed BRDF ρ' after computation of the weights. A common compression function is the logarithmic function, in which case Equation 2 becomes a homomorphic factorization.

2.2 Natural Lighting

Prior work relied on point sample measurements of ρ for a set of incoming-outgoing direction pairs to estimate the weights w. In contrast, in this paper we aim to estimate the weights w from an observation under natural lighting. Assuming the lighting L is distant (i.e., it only depends on the incident direction $\omega_i = (\phi_i, \theta_i)$), and ignoring interreflections, we can formulate the observed radiance y as:

$$y(\omega_o) = \int_\Omega \rho(\omega_i, \omega_o) \cos(\theta_i) L(\omega_i) d\omega_i, \quad (3)$$

where $\cos(\theta_i)$ is the foreshortening, and Ω is the upper hemisphere of incident directions. Due to linearity of light transport, we can express Equation 3 in terms of corresponding basis observations y:

$$y = Yw, \qquad (4)$$

where the weights w are the same as in Equation 1, and thus can be used to reconstruct ρ. The basis images $Y = [y_0, ..., y_n]$ are the observations of the measured basis BRDFs b_i under the same conditions:

$$y_i = \int_\Omega b_i(\omega_i, \omega_o) \cos(\theta_i) L(\omega_i) d\omega_i. \quad (5)$$

3 BRDF Likelihood Modeling

3.0 Gaussian Mixture Model

We propose to model the likelihood of BRDFs by a Gaussian mixture model (GMM):

$$P(\rho) = \sum_{j=1}^{k} \pi_j \mathcal{N}(\rho|\mu_j, \Sigma_j), \qquad (6)$$

where π_j are the mixing coefficients of the j-th normal distribution $\mathcal{N}$ with mean μ_j and covariance matrix Σ_j.

3.1 Expectation-Maximization

An effective method for computing the parameters $\Theta = (\pi, \mu, \Sigma)$ is the Expectation Maximization algorithm using the MERL BRDFs b_i as observations. For this we define a latent variable $\gamma_j(b_i)$ that indicates the likelihood of the j-th Gaussian given a MERL BRDF b_i:

$$\gamma_j(b_i) = P(j|b_i), \qquad (7)$$

$$= \frac{P(j)P(b_i|j)}{P(b_i)}, \qquad (8)$$

$$= \frac{\pi_j \mathcal{N}(b_i|\mu_j, \Sigma_j)}{\sum_{j=1}^{k} \pi_j \mathcal{N}(b_i|\mu_j, \Sigma_j)}. \qquad (9)$$

Expectation minimization iterates between estimating the latent variable $\gamma_j(b_i)$ (E-step, Equation 9), and the model parameters (M-step):

$$\pi_j = \frac{1}{n} \sum_{i}^{n} \gamma_j(b_i), \qquad (10)$$

$$\mu_j = \frac{\sum_{i}^{n} \gamma_j(b_i) b_i}{\pi_j}, \qquad (11)$$

$$\Sigma_j = \frac{\sum_{i}^{n} \gamma_j(b_i)(b_i - \mu_j)(b_i - \mu_j)^T}{\pi_j}. \quad (12)$$

We iterate until the log-likelihood over the MERL BRDFs converges:

$$\log P(B|\Theta) = \sum_{i}^{n} \log \sum_{j}^{k} \pi_j \mathcal{N}(\rho|\mu_j, \Sigma_j). \quad (13)$$

To bootstrap the EM algorithm, we perform a standard k-mean clustering, and initialize π_j as the ratio of assigned BRDFs to the j-th cluster over the total number of MERL BRDFs (i.e., n).

3.1 Curse of Dimensionality

A practical problem is that the number of observations n is significantly lower than the dimensionality of the space (i.e., p).We therefore apply singular value decomposition (SVD) to express the observations in a n dimensional space U:

$$B = USV^T. \qquad (14)$$

However, this is still a 300 dimensional space. A key issue is that even for a moderate number of dimensions any distance is very large, and thus the distance to the means μ_j are large too. Consequently, the likelihood of each Gaussian mixture (Equation 9) will always be very low and it can potentially cause numerical instabilities. To resolve this issue, we perform expectation maximization in a reduced space, and only keep the coefficients belonging to the N largest singular values. In other words, we perform expectation maximization (i.e., soft clustering) on a projection to an N dimensional space, and approximate the likelihood: $P(\rho) \approx P(\hat{U}^T\rho)$, where $\hat{U}$ is the N dimensional basis (i.e., the first N vectors in U).

3.2 Discussion

We found that $N = 4$ offers a good balance between accuracy and numerical stability. A second parameter that needs to be set is the number of Gaussian mixtures K. If the number of Gaussians is too low, then $P(\hat{U}^T\rho)$ only offers a coarse approximation. In practice we found that $K = 4$ offers a good approximation that nicely categorizes the materials in four recognizable distinct material classes: *"diffuse and glossy"* materials (137 materials), *"plastics/phenolics"* (99 materials), *"metals"* (24),

and *"specular plastics/paints"* (40 materials); we determine membership to a material class by assigning the material to the material class with the maximum $\gamma_j(b_i)$ likelihood. Figure 1 shows a plot of a 2D multi-dimensional scaling of the 4D projected coordinates of the MERL BRDFs, as well as a color-coding to indicate for which material class the material has the highest affinity. Note that even though the diffuse-like material class contains 137 materials, the multi-dimensional scaling places them all close together.

4 Data-driven Model Estimation

4.1 MAP Estimation

We express the likelihood of the observation given a BRDF as:

$$P(y|\rho) = \mathcal{N}(Yw - y|\mu, \sigma), \qquad (15)$$

where μ and Σ is the expected mean error and standard deviation on the reconstructions. We assume that the mean error is close to zero ($\mu = 0$), and σ is proportional to the expected measurement error (e.g., camera noise).
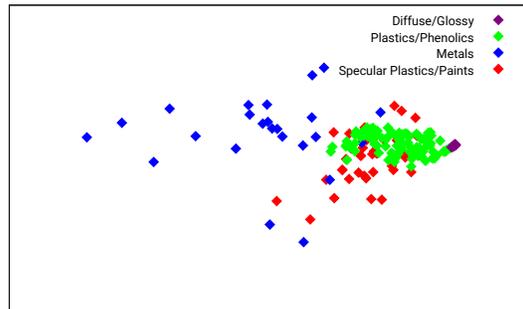

Figure 1: 2D multi-dimensional scaling of the projected MERL BRDFs $\hat{U}^t B$ and a color-coding of the respective material classes derived from the 4D approximation of the BRDF likelihood modeled by a Gaussian mixture model.

Given the likelihood $P(\hat{U}^T \rho)$ expressed by the Gaussian mixture model (Equation 6), we can then formulate the MAP estimation (**??**) as:

$$\underset{w}{\mathrm{argmin}} \left(\frac{||Yw - y||^2}{\sigma^2} + \log \sum_j \pi_j \mathcal{N}(\hat{U}^T Bw | \mu_j, \Sigma_j) \right) \tag{16}$$

The first term is the data term that indicates how well the BRDF $\rho = Bw$ can explain the observation y, and the second term indicates how plausible the reconstructed BRDF ρ (projected in the 4 dimensional space $\hat{U}$) is.

However, directly solving for the BRDF weights w using Equation 16 is not practical because of two key practical issues:

1. *Non-linear:* Equation 16 is highly non-linear due to the sum of the log-likelihoods in the second term, and which is difficult to optimize.

2. *Gaussian Mixture Model Accuracy for $P(\rho) \approx P(\hat{U}^T \rho)$:* We approximated the likelihood of the BRDF by a four dimensional Gaussian mixture model. This reduction in dimensionality was necessary due to the curse of dimensions. However, it also makes an implicit assumption, namely that the BRDF lies not too far from the space of plausible BRDFs. Since the likelihood is only determined based on 4 dimensions (and thus only regularizes these four), the other 296 dimensions can be set to any value (including unreasonable values that result in an implausible BRDF).

4.2 Linear MAP Approximation

To alleviate the above two practical issues, we exploit the observation that the likelihood of a basis BRDF b_i belonging to a material class m is for most basis BRDFs equivalent to an indicator function:

$$\gamma_j(b_i) \approx \delta_{i,m}. \tag{17}$$

This implies that the overlap between the Gaussians in the Gaussian mixture model is limited. Armed with this observation, we therefore propose to compute a candidate BRDF for each material class $j \in [1, k]$:

$$\underset{w^{(j)}}{\mathrm{argmin}} \left(\log P(y|\rho, j) + \log P(\rho|j) \right). \tag{18}$$

Given the set of candidate solutions $w' = \{w^{(1)}, .., w^{(k)}\}$, we then pick the best candidate that best reconstructs the BRDF.

By a-priori assuming that a BRDF belongs to a material class j, it is possible that there is a significant mismatch between the target material and the material class. For example, attempting to model a mirror-like specular material using the diffuse material class is unlikely to produce a satisfactory result. Consequently, we cannot simply rely on the likelihood $P(\hat{U}^T \rho)$ based on the 4 dimensional Gaussian mixture model. We will therefore further exploit the observation of the limited overlap of the Gaussians in the mixture model, and approximate the solution per material class by enforcing that it lies in the convex hull of the subspace spanned by the BRDFs assigned to the material class, and only rely on Equation 15 to pick the best candidate from w'. We ignore the standard deviation (i.e., $\sigma = 1$) in Equation 15 as it only acts as a scale (in the log-likelihood) that does not affect the selection of the best reconstruction (i.e., minimum log-likelihood).

4.3 Per-Material Class Linear Data Term

We define the data-term similarly as in the general non-linear case, except that we only use the basis BRDFs that belong to the same material class:

$$\log P(y|\rho, j) = ||Y^{(j)} w^{(j)} - y||^2, \tag{19}$$

where $Y^{(j)}$ is the set of observations that correspond to the basis BRDFs assigned to the j-th

material class (i.e., the materials b_i for which $\gamma_j(b_i)$ is maximal).

4.4 Per-Material Class Linear Likelihood Term

We express the per-material class likelihood by a single Gaussian model. We directly compute this probability on the BRDF weights $w^{(j)}$:

$$P(\rho|j) = \mathcal{N}(w^{(j)}, \mu'_j, \Sigma'_j), \qquad (20)$$

where: $\mu'_j = \frac{1}{c_j}$, and c_j is the number of basis BRDFs in the j-th material class. Note that $Y^{(j)}\mu'_j$ is equivalent to the mean of the BRDFs in the material class.

4.5 Linear Least Squares Estimation

Both Equation 19 and (the log likelihood of) Equation 20 are quadratic terms that define a linear system in terms of w that can be solved using a regular linear least squares. However, both terms can have a vastly different magnitude. The magnitude of the data-term depends on the error on the rendered image of the estimated BRDF. This image error depends on the resolution, the overall intensity of the lighting, and the reflectivity of the material. Similarly, the magnitude of the likelihood term depends on the number of basis BRDFs per material class. We therefore add a balancing term:

$$\lambda_j = \frac{\lambda ||y||^2}{c_j}, \qquad (21)$$

where $||y||^2$ is the total squared pixel intensities in the observation. We expect that the overall intensity of the observation is directly proportional to the lighting intensity and reflectivity of the BRDF, and hence the overall scale of the image error. λ is a user set constant that depends on the qualities of the lighting. An ill-conditioned lighting condition requires a larger λ value (e.g., a low frequency lighting environment is ill-conditioned for estimating specular properties[9]). In practice we found

that $\lambda = 0.5$ works well for many lighting environments, and forms a good starting point for fine-tuning λ.

The final linear least squares is:

$$\underset{w^{(j)}}{\mathrm{argmin}} \left(||Y^{(j)}w^{(j)} - y||^2 + \lambda_j \frac{||w^{(j)} - \mu'_j||^2}{\Sigma_j^2} \right) \quad (22)$$

4.6 Color

Our discussion until now only considered monochrome BRDFs; we used all color channels from the MERL BRDFs as separate basis BRDFs. A straightforward strategy for estimating a non-monochrome BRDF with three color channels, would be to execute the estimation separately for each color channel, and combine the three reconstructed monochrome BRDF into a single RGB BRDF. However, it is possible that a solution from a different material classes j is selected for each of the three color channels. Because the set of basis BRDFs for each material class are disjunct, there can be slight differences in the constructed BRDF shape for each color channel, which in turn can result in color artifacts in the combined BRDF. We circumvent this potential problem by combining the three color channels after obtaining the candidate BRDFs, and performing the selection on the RGB BRDF instead of each color channel separately. Hence, each color channel will be reconstructed with the same set of basis BRDFs.

4 Results

4.6 Experiment Setup

We demonstrate our method on simulated captures in order to fully control all parameters. We generate "observations" under natural lighting, by rendering a sphere lit by a light probe[1] using Mitsuba.[2] All generated images are radiometrically linear, and we only tone map

them for display. All results shown in this paper were tone mapped by a simple gamma 2.2 correction and a virtual exposure (i.e., scale factor) of 1.0; all pixel values above 1.0 and below 0.0 are clipped to the respective clipping values. We use the BRDFs in the MERL database[3] for generating observations. For each MERL BRDF, we compute a novel Gaussian mixture model on the 297 remaining MERL BRDFs, and only use these 297 MERL BRDFs for reconstruction. As noted in the prior sections, we compute the Gaussian mixture model on a $N = 4$ dimensional reduced space, and use $K = 4$ Gaussians in the mixture model. All reconstructions are generated with a fixed balancing factor $\lambda = 0.5$.

4.6 Reconstruction Results

Figure 2 shows reconstructions of 7 selected materials under two different light probes (i.e., *Eucalyptus Grove* and *Galileo's Tomb*). For each reconstruction (and the reference), we show a visualization of the reference/reconstructed BRDF under a natural lighting condition (i.e., *Uffizi Gallery*; different than the lighting condition under which the BRDF was reconstructed) and a directional light (i.e., a slice of the BRDF for a single incident direction for all outgoing directions). These results show that our method is able to reconstruct plausible BRDFs for a wide range of materials from a single photograph of a spherical object under natural lighting.

4.6 Per-Material Class Reconstruction

Figure 3 illustrates, for a selection of 4 materials, reconstructed under the *Uffizi Gallery* light probe, that the reconstructions per material class are different, and that depending on the material a different class' reconstruction is selected. We show a visualization of the reference BRDF and the reconstructions per cluster under a natural lighting condition (i.e., *Euca-*

lyptus Grove) and a directional light. We also list the log-likelihood of the observation given the BRDF (Equation 15) below each cluster, and mark the final selected solution (i.e., minimum). For reference, we also show the linear least squares solution: $\text{argmin}_w ||Yw - y||^2$. As expected this yields the lowest reconstruction error (since it explicitly optimizes for this). However, the linear least squares solution does not always yield a plausible result when visualized under a different lighting condition. This is not only clearly visible under the directional light source, but also under other natural lighting conditions other than the original observed lighting (e.g., the black spot in the center of the visualizations under the *Eucalyptus Grove* light probe for *Steel* and *Red Metallic Paint*).

4.7 Discussion

While our selection criterion does in the majority of cases select the best reconstruction from the different material classes, we found that in a very few cases cases it does not select the best reconstruction, and a better reconstruction can be observed in a different material class. We currently used a λ balancing factor of 0.5 for all our reconstructions. This λ is a compromise to produce the best result over all materials. Despite the material class and scene dependent scale factor (Equation 21), we observe that this lambda terms tends to affect the *"diffuse"* and *"plastics / phenolics"* stronger, and the *"metals"* and *"specular plastics / paints"* less. These latter two material classes exhibit not only a lower number of materials (for which we compensate), but we can also observe in Figure 1 that they are also spread out further. Consequently, the density of these material classes is significantly lower. We have demonstrated our method on spherical shapes. For other convex shapes that cover the full range of normal directions, we can remap the pixels based on their underlying normal to a sphere and apply the same reconstruction algorithm. Alterna-

tively, we can also render Y on the same shape as the observations and directly solve the linear least squares.

5 Conclusion

In this paper we presented a novel method for estimating the parameters of a fully linear data-driven BRDF model from just a single photograph of a object with known shape under un-controlled, but known, natural lighting. Our estimation method does not require any non-linear optimization, and only requires solving 4 linear least squares problems. Our method requires modest precomputations: a Gaussian mixture model clustering for the basis BRDFs, and for each natural lighting conditions, renderings of each basis material. We demonstrated the accuracy and robustness of our method on the MERL BRDF database.

For future work we would like to explore better selection criteria and a per-material class λ_j density correction factor.

5 References

[1] P. Debevec. Light probe gallery. `http://www.pauldebevec.com/Probes/`, 1998.

[2] W. Jakob. Mitsuba: Physically based renderer. `https://www.mitsuba-renderer.org`, 2010.

[3] Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillan. A data-driven reflectance model. *ACM Trans. Graph.*, 22(3):759–769, July 2003.

[4] Wojciech Matusik, Hanspeter Pfister, Matthew Brand, and Leonard McMillan. Efficient isotropic BRDF measurement. In *Rendering Techniques*, pages 241–248, 2003.

[5] NASA-Directorates. Mission directorates education and outreach. `http://www.nasa.gov/offices/education/missions/index.html`. Accessed: 2016-30-16.

[6] NASA-HumanExploration. Human exploration and operations: Human exploration in and beyond low-earth orbit. `http://www.nasa.gov/directorates/heo/education/index.html#.VrNmx7IrK00`. Accessed: 2016-30-16.

[7] NASA-Science. Nasa science: About us. `http://science.nasa.gov/about-us/`. Accessed: 2016-30-16.

[8] Jannik Boll Nielsen, Henrik Wann Jensen, and Ravi Ramamoorthi. On optimal, minimal brdf sampling for reflectance acquisition. *ACM Trans. Graph.*, 34(6), October 2015.

[9] Ravi Ramamoorthi and Pat Hanrahan. A signal-processing framework for inverse rendering. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, pages 117–128, 2001.

[10] Neal E Seymour, Anthony G Gallagher, Sanziana A Roman, Michael K O'Brien, Vipin K Bansal, Dana K Andersen, and Richard M Satava. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of surgery*, 236(4):458–464, 2002.

[11] Zexiang Xu, Jannik Boll Nielsen, Jiyang Yu, Henrik Wann Jensen, and Ravi Ramamoorthi. Minimal brdf sampling for two-shot near-field reflectance acquisition. *ACM Trans. Graph.*, 35(6):188:1–188:12, November 2016.
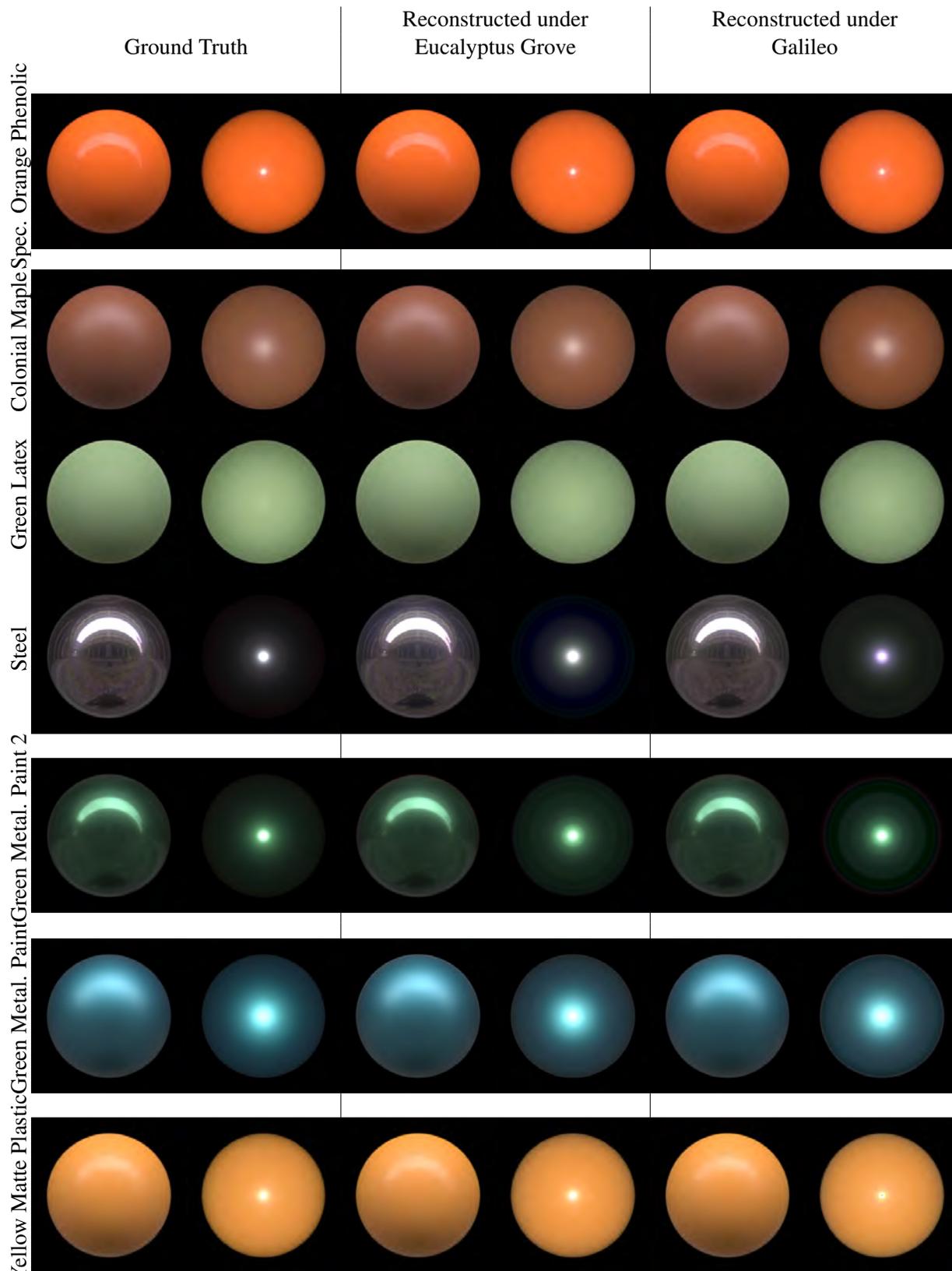
Figure 2: Data-driven BRDF reconstructions from a single photograph of a spherical object captured under the *Eucalyptus Grove* and the *Galileo's Tomb* light probe. We visualize the reference and reconstructed BRDFs under the *Uffizi Gallery* light probe and a directional light.
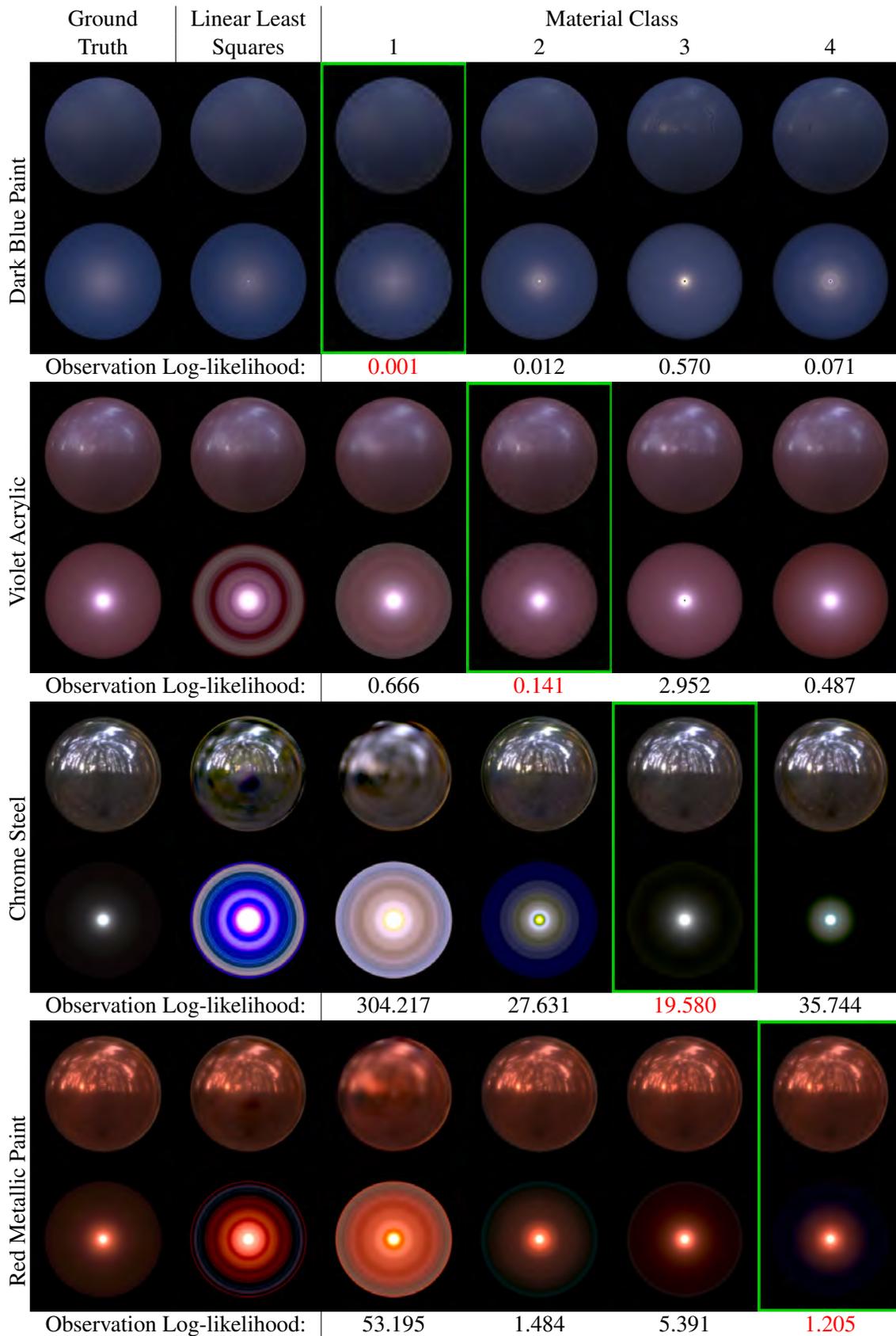
Figure 3: Reconstructions for each material class for 4 selected materials under the *Uffizi Gallery* light probe, and revisualized with the *Eucalyptus Grove* light probe and directional lighting.

Cooper