Workshop on Material Appearance Modeling (2020)
H. Rushmeier and R. Klein (Editors)

An Adaptive Metric for BRDF Appearance Matching

J. Bieron P. Peers

College of William & Mary

Abstract

Image-based BRDF matching is a special case of inverse rendering, where the parameters of a BRDF model are optimized based on a photograph of a homogeneous material under natural lighting. Using a perceptual image metric, directly optimizing the difference between a rendering and a reference image can provide a close visual match between the model and reference material. However, perceptual image metrics rely on image-features and thus require full resolution renderings that can be costly to produce especially when embedded in a non-linear search procedure for the optimal BRDF parameters. Using a pixel-based metric, such as the squared difference, can approximate the image error from a small subset of pixels. Unfortunately, pixel-based metrics are often a poor approximation of human perception of the material's appearance. We show that comparable quality results to a perceptual metric can be obtained using an adaptive pixel-based metric that is optimized based on the appearance similarity of the material. As the core of our adaptive metric is pixel-based, our method is amendable to image-subsampling, thereby greatly reducing the computational cost.

1. Introduction

Matching the appearance of an isotropic homogeneous material with an analytic BRDF model is not straightforward. It is extremely unlikely that there exists a set of parameter values that will cause the model to behave *exactly* like the real reflection function of an object. Even if the material's BRDF has been exhaustively measured, directly minimizing the difference (according to some predetermined cost function) between measured BRDF samples and the model's predictions does not always produce a perceptually close match [BP20]. Furthermore, it may be the case that the only available information about the BRDF of a material is one or more photographs of an object composed of the material under natural lighting. If scene properties such as object shape and environment lighting are known, then the process of reproducing the material's appearance with a BRDF model becomes an inverse rendering problem where the parameters of the appearance model are optimized to minimize the difference between reference images of the material and corresponding synthesized images using the analytic model. Using a perceptual image-difference metric instead of a simple squared error typically results in a more faithful reproduction of the visual appearance of the material. Matching the appearance from a photograph based on a perceptual metric is computationally expensive because a perceptual metric gauges the image quality based on image features. Consequently, in each step of the non-linear search algorithm (for the optimal BRDF parameters), the entire scene must be rendered. Furthermore, perceptual image metrics often result in a more difficult to navigate error-landscape.

In this paper we investigate a two-stage adaptive metric for BRDF matching inspired by [BP20]. In our first stage, we will use a pixel-wise adaptive metric to generate candidate BRDF parameters by inverse rendering. In the second stage, we use a perceptual metric to select the best candidate based on the visual difference as predicted by the perceptual image difference metric. We will show that the first stage can be greatly accelerated by approximating the pixel-wise metric on a small subset of pixels without adversely affecting the quality of the BRDF match. Furthermore, our experiments indicate that it is essential to adapt pixel-based metrics to the materials and the scene, giving clue to why an L_2 metric alone is not adequate for accurate BRDF matching.

2. Related Work

Inverse Rendering BRDF matching can be seen as a specialized application of inverse rendering which optimizes any subset of scene parameters such that the difference between the reference photographs and corresponding renderings is minimized [Mar98]. We refer to [PP03] for a general overview of inverse rendering methods, and to [WK15] and [DRS08] for an overview geared towards appearance modeling. The majority of inverse rendering methods rely on a squared image difference metric in order to use fast linear or non-linear optimization methods often accompanied with hand-crafted regularization terms to local minima to avoid ambiguities in the solution space (e.g. [DCP*14, LN16, BM12]). However, some researchers have used different image difference metrics. For example, Khungurn et al. [KSZ*15] avoid the need for exact per-pixel matching of photographs of fabric and renderings when matching the parameters of micro-appearance models

DOI: 10.2312/mam.20201137

by comparing the error on the average of the images. Another example is Shacked and Liskinski's [SL01] lighting design by inverse rendering with a custom perceptual error metric.

CSSIM We have opted to use *Color Structural Similarity Metric (CSSIM)* [LPU*13] as the whole-image metric due to its proven effectiveness for appearance modeling [HFM16, BP20].

Variable Error Metric Bieron and Peers [BP20] introduced a novel adaptive BRDF fitting metric for reflectance data. They propose a two stage approach where in the first stage an adaptive metric with a free parameter γ is used to fit BRDF parameters to match densely sampled reflectance measurements, generating candidate BRDF fits. In the second stage, the best BRDF fit from among the candidates is selected according to the perceptual difference between renderings of the measured reflectance and the BRDF fit. Building on this approach, we consider the problem of fitting a BRDF model to a target image through inverse rendering without access to the underlying reflectance of the target material.

3. Method

Cost Function We desire a pixel-wise image-to-image comparison metric that does not rely on image structure in any way (i.e., the error for each pixel is independent of all neighboring pixels). Thus, we want an cost function C that can be expressed in terms of the difference between corresponding pixels p in A and B. For this we borrow the compression function used in [BP20] and apply it to images. This gives us

$$C(A,B,\gamma) = \sum_{p \in A,B} (A[p]^{\frac{1}{\gamma}} - B[p]^{\frac{1}{\gamma}})^2 \qquad (1)$$

which can also be seen as the squared difference of tonemapped images without clamping to $[0,1]$. It is not the case that the best BRDF fit corresponds with the tonemapping function used for displaying the images. As in [BP20] the metric tends to give BRDF matches with sharper specular highlights and brighter (often miscolored) diffuse reflectance for low values of γ. For higher γ values, the BRDF matches exhibit more color fidelity but blurry highlights.

Two-Stage Approach The first question we must answer is whether or not for some γ our cost function $C(A,B)$ will give a BRDF match which is close to the result of directly optimizing a perceptual metric $P(A,B)$, where P is CSSIM in our case. We observe that for some γ optimizing to minimize C gives a good match to directly minimizing P (we use Matlab's *patternsearch* for both). Figure 1 shows the results of minimizing CSSIM directly compared with the choosing BRDF match with an optimal γ.

Secondly, we examine whether the γ that leads to the lowest CSSIM error varies between materials. We find that it does. Figure 2 shows the relationship between γ and the CSSIM error for three materials.

The above observation yields the following potential two-stage algorithm: first, for each γ value in some range (i.e., $[1.0, 1.1, \cdots, 2.9, 3.0]$ in our implementation), optimize the BRDF parameters with C as the cost function. Then, choose the resulting BRDF match with the lowest CSSIM error from among all the γ's checked. The computational cost of this algorithm exceeds directly

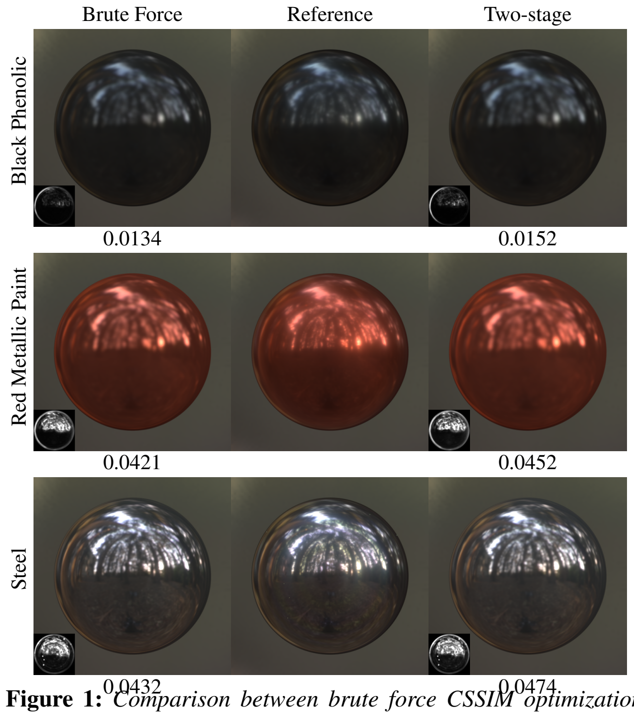

Figure 1: *Comparison between brute force CSSIM optimization and our two-stage approach using* Eucalyptus Grove *lighting. While the two-stage approach does not achieve the optimal CSSIM error, the error and appearance are close.*

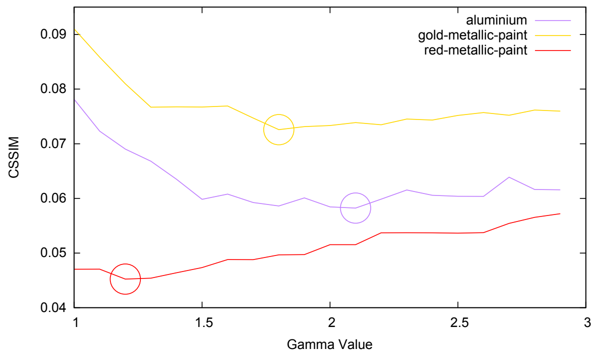

Figure 2: *CSSIM error versus γ value plotted for* Gold Metallic Paint 3, Aluminium, *and* Red Metallic Paint. *The minimum CSSIM error is circled for each graph.*

using CSSIM as our cost function because C requires the same rendering costs as optimizing with CSSIM directly, but the pixel-wise nature of C presents a potential solution.

Subsampling We find that performing our first-stage optimizations of C over a subset of the pixels in A and B performs similarly to using all the pixels in the images. This makes our strategy practical, as we can drastically reduce the computation cost of our 21 non-linear optimizations with C by rendering only a fraction of the pixels, and then we render the all the BRDF candidates at full resolution and use CSSIM to select the best match.

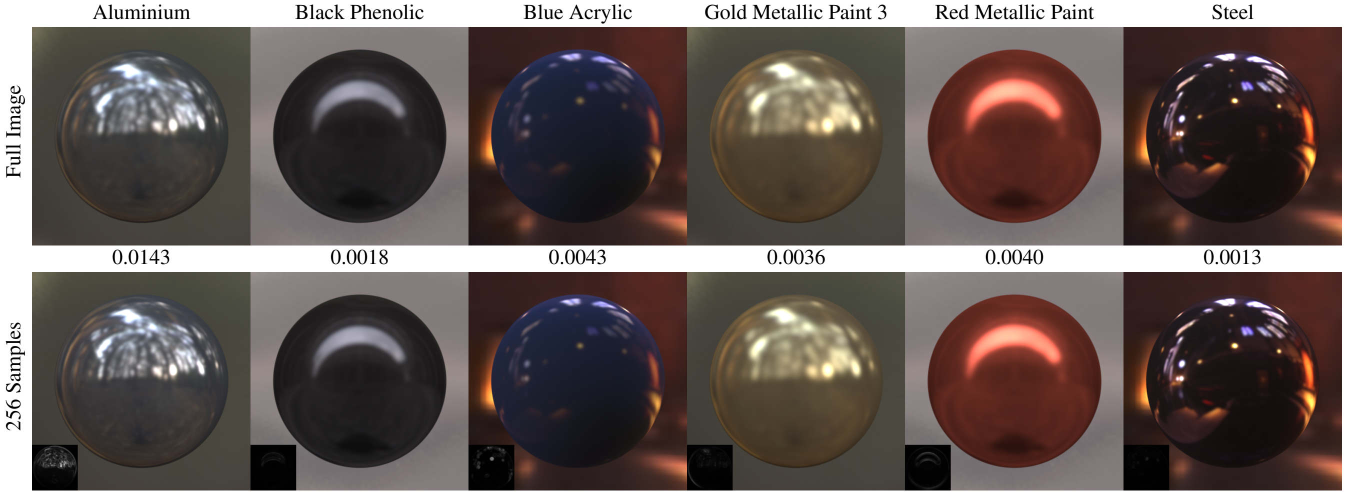

Figure 3: *Comparison between our two-stage inverse rendering approach executed on the full image (top row) versus only 256 selected pixels (bottom row) on a selection of* 6 *materials performed under the lighting shown. The difference image (inset) and CSSIM error (middle row) are between the full image solution and the corresponding importance sampled solution.*

Table 1: *Average CSSIM error of our two-stage inverse-rendering results over* 6 *selected materials (Figure 3) under three different lighting conditions and for varying number of samples.*

	32	64	256	Full Image
Grace Cathedral	0.0735	0.0656	0.0631	0.6282
Uffizi Gallery	0.0684	0.0579	0.0532	0.6115
Eucalyptus Grove	0.0449	0.0410	0.0410	0.0436

We generate our subset of pixels once per scene, and reuse it for all first-stage optimizations. We use a Monte Carlo importance sampling strategy: we select pixel positions proportionally to the pixel's intensity in the reference photograph, and weight each pixels' contribution to the error by one over the PDF of choosing that pixel, to avoid biasing. The key idea is that errors are most likely to be significant in bright areas. Figure 3 compares our BRDF fitting results on 6 selected materials with 256 samples compared to a full image sampling. The inverse rendering, as well as visualization, was performed under the lighting used in the figure.

The speed-up of this subsampling scheme is directly proportional to how many pixels we need to render. The ideal number of samples depends on the scene properties and lighting. We found that 256 pixels was a "safe" choice. However, we empirically found that even at only 32 samples good results can still be obtained. Table 1 summarizes the average CSSIM error of the BRDF fits obtained with our subsampling technique over the 6 materials in Figure 3 under three different lighting conditions.

With 256 pixels this technique gives a factor 10 speed-up compared with directly optimizing CSSIM. The cost of our first stage is 21 optimizations at only $\frac{1}{256}$ of the rendering cost. The fixed costs of 21 full resolution renders and CSSIM on those images are negligible.

4. Discussion

We validate our two-stage adaptive metric for BRDF matching using synthetic target photographs of materials from the MIT-MERL

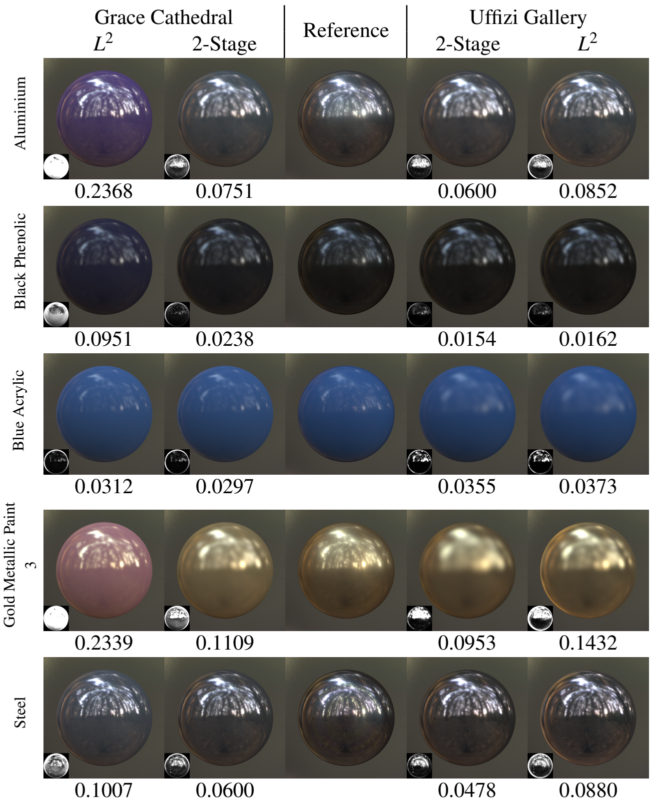

Figure 4: *The adverse effect of suboptimal lighting demonstrated on our method (with* 256 *samples) compared with L_2 under the* Uffizi Gallery *and* Grace Cathedral *lighting.*

BRDF database [MPBM03] of spherical shapes under known natural lighting. Figure 4 shows results for our method for five materials matched under two lighting conditions and visualized under a third. Note that our two-stage approach not only outperforms a naive L_2 based approach, but it is also considerably faster. We observe that while the lighting under which the inverse rendering is performed has a clear impact on the BRDF match, our method is much more robust to changes in lighting.

We observed that the optimal γ value depends not only on the material and lighting as in [BP20] but also on the subset of pixels selected for our stage one BRDF fits. For 32 pixel rendering under the *Grace Cathedral* light probe the effect was especially noticeable. This remained true when we added a clamping to a $[0,1]$ range, though the visual impact was lessened. From this we conclude that the second stage of our algorithm helps regularize the solutions from our first stage which may be biased towards more specular or diffuse dominated solutions depending on the particular pixels chosen.

Reference	64×64	128×128	256×256	2-Stage

	0.0281	0.0209	0.0182	0.0225

Figure 5: *Using low resolution images to speed up brute force CSSIM optimization produces less sharp BRDFs. Note that the 2-Stage result used only 32 pixels in the first stage, yet it outperforms the brute force optimization with* $64\times64 = 4096$ *pixels.*

We also tested the impact of speeding up the inverse rendering process by reducing the resolution of our scene. Figure 5 shows the results of brute force CSSIM optimization at different resolutions, all rendered at the same resolution. Sampling a small fraction of the image to fit using our variable metric and then choosing the best BRDF match at full resolution proved not only faster than fitting at low resolution, it also gave more accurate results.

5. Conclusion

The method described in [BP20] can extend to appearance matching, and we find that the optimal γ varies with material and scene properties. We found that using even a small sampling of pixels in an HDR image, results of similar quality to directly optimizing the full image with a perceptual metric can be obtained in only a fraction of the time. This perhaps implies that more measurements are not always the answer to better BRDF acquisition; the metric(s) used should also be considered. Using less data with multiple metrics might improve BRDF capture in the wild.

References

[BM12] BARRON J. T., MALIK J.: Shape, albedo, and illumination from a single image of an unknown object. *CVPR* (2012). 1

[BP20] BIERON J., PEERS P.: An adaptive brdf fitting metric. *Eurographics Symposium on Rendering 2020* (2020). 1, 2, 4

[DCP*14] DONG Y., CHEN G., PEERS P., ZHANG J., TONG X.: Appearance-from-motion: Recovering spatially varying surface reflectance under unknown lighting. *ACM Trans. Graph. 33*, 6 (2014), 193:1–193:12. 1

[DRS08] DORSEY J., RUSHMEIER H., SILLION F.: *Digital Modeling of Material Appearance*. Morgan Kaufmann Publishers Inc., 2008. 1

[HFM16] HAVRAN V., FILIP J., MYSZKOWSKI K.: Perceptually motivated brdf comparison using single image. *Computer Graphics Forum 35*, 4 (2016). 2

[KSZ*15] KHUNGURN P., SCHROEDER D., ZHAO S., BALA K., MARSCHNER S.: Matching real fabrics with micro-appearance models. *ACM Trans. Graph. 35*, 1 (Dec. 2015), 1:1–1:26. 1

[LN16] LOMBARDI S., NISHINO K.: Reflectance and illumination recovery in the wild. *IEEE PAMI 38*, 1 (2016), 129–141. 1

[LPU*13] LISSNER I., PREISS J., URBAN P., LICHTENAUER M. S., ZOLLIKER P.: Image-difference prediction: From grayscale to color. *IEEE TIP 22*, 2 (2013), 435–446. 2

[Mar98] MARSCHNER S.: *Inverse Rendering for Computer Graphics*. PhD thesis, Cornell University, 1998. 1

[MPBM03] MATUSIK W., PFISTER H., BRAND M., MCMILLAN L.: A data-driven reflectance model. *ACM Trans. Graph. 22*, 3 (2003), 759–769. 3

[PP03] PATOW G., PUEYO X.: A Survey of Inverse Rendering Problems. *Computer Graphics Forum* (2003). 1

[SL01] SHACKED R., LISCHINSKI D.: Automatic Lighting Design using a Perceptual Quality Metric. *Computer Graphics Forum* (2001). 2

[WK15] WEINMANN M., KLEIN R.: Advances in geometry and reflectance acquisition (course notes). In *SIGGRAPH Asia 2015 Courses* (2015), pp. 1:1–1:71. 1