Eurographics Symposium on Rendering 2019
T. Boubekeur and P. Sen
(Guest Editors)

Volume 38 (2019), Number 4

Learned Fitting of Spatially Varying BRDFs - Supplementary Material

S. Merzbach[1], M. Hermann[2], M. Rump[1,2] and R. Klein[1]

[1]University of Bonn, Institute of Computer Science II, Germany
[2]X-Rite Inc.

1. Acquisition Details

Using Pantora's anisotropic textile preset, the overall acquisition is performed by rotating the material at five 45° steps, while capturing images with all four cameras. For each captured image, individual LED point lights (color filtered or not) are switched on. The entire process results in 348 (4 cameras × 3 rotation steps × 29 LEDs) point lit panchromatic images. 100 of those point lit images have additional color information through band filtered illumination available. Finally, the device captures images for each turntable rotation, viewed by each camera, while moving the linear light source. Depending on a glossiness preset, the linear light source declination angle step size is 4° (#steps = 14) for low, 2° (#steps = 28) for medium and 0.5° (#steps = 165) for high gloss materials. This results in (#steps × 5 × 4) ∈ {280, 560, 3300} additional images for low, medium or high gloss materials respectively.

2. Postprocessing

The raw measurement images are postprocessed by the Pantora software. All raw photographs are combined into radiometrically calibrated high dynamic range (HDR) images. Geometric Aruco [GMMM14] markers are detected around the sampleholder to calibrate the turntable rotation and determine the exact 3D position of each camera with respect to the sample. Then the structured light images are decoded to find 3D correspondences for all pixels. Those are interpolated into a heightmap, which has pixels in a reference coordinate system defined by the top view camera. In the next steps, all measured images (panchromatic, color and linear light source) are projected from the respective camera's point of view onto the 3D surface imposed by the heightmap in the top view reference coordinate system. Together with the heightmap, those radiometrically calibrated HDR images with pixel-wise correspondences are the inputs to the final stage of the postprocessing pipeline, the SVBRDF fitting.

3. SVBRDF Model

To be able to describe the model in detail, we first need to define some preliminaries. We will specify the Ward BRDF for a point $\mathbf{p}$ on the surface. For SVBRDFs, $\mathbf{p}$ is usually expressed in pixel coordinates x and y, and thus, the following parameter maps can be considered as textures. The heightmap, $H(\mathbf{p})$, is used to compute a geometric tangent frame, defined by geometric normal, tangent and bitangent: $\{\mathbf{n}_g(\mathbf{p}), \mathbf{t}_g(\mathbf{p}), \mathbf{b}_g(\mathbf{p})\}$. The derived geometric normals provide only a rather coarse resolution, therefore the SVBRDF is evaluated with a shading normal, which is stored in an additional normal map $\mathbf{n}_s(\mathbf{p})$. This shading normal replaces $\mathbf{n}_g(\mathbf{p})$ in the aforementioned coordinate frames. To ensure a valid basis, tangent and bitangent need to be re-orthogonalized by a Gram-Schmidt step. Finally, the tangent frames are rotated around $\mathbf{n}_s(\mathbf{p})$ by an angle $\alpha(\mathbf{p})$, stored in an anisotropy map. The resulting three basis vectors form an orthonormal basis for each pixel in the SVBRDF. To evaluate the BRDF model per pixel, light and view directions need to be transformed from the global to these local coordinate systems. We define the local directions as $\mathbf{l}, \mathbf{v} \in \mathbb{R}^3$. First, $\mathbf{l}$ and $\mathbf{v}$ are used to compute an unnormalized halfway vector $\mathbf{h} = \mathbf{l} + \mathbf{v}$. The normalized version $\mathbf{h}' = {}^{\mathbf{h}}/_{\|\mathbf{h}\|}$ is necessary for the Fresnel term. The anisotropic Ward [War*92; GD10] model with Fresnel reflection term can then be defined as follows:

$$f(\mathbf{p}, \mathbf{h}, \mathbf{h}') = \frac{a_d(\mathbf{p})}{\pi} + \frac{a_s(\mathbf{p}) F(F_0(\mathbf{p}), \langle \mathbf{v}, \mathbf{h}' \rangle)}{\pi \sigma_x(\mathbf{p}) \sigma_y(\mathbf{p}) h_z^4} e^d,$$

with exponent

$$d = -\frac{\frac{h_x^2}{\sigma_x(\mathbf{p})^2} + \frac{h_y^2}{\sigma_y(\mathbf{p})^2}}{h_z^2},$$

where $\mathbf{a}_d(\mathbf{p})$ is the diffuse albedo, $\mathbf{a}_s(\mathbf{p})$ the specular albedo, $F_0(\mathbf{p})$ the Fresnel reflectance at 0° incidence. The anisotropic specular reflection lobes are parameterized by $\sigma_x(\mathbf{p})$ and $\sigma_y(\mathbf{p})$. The Schlick Fresnel approximation [Sch94] is defined as

$$F(F_0(\mathbf{p}), \theta) = F_0(\mathbf{p}) (1 - F_0(\mathbf{p})) (1 - \cos\theta)^5.$$

Note that contrary to the traditional Ward model and its variants, the specular albedo $\mathbf{a}_s(\mathbf{p})$ is no longer bounded by 1 but instead its product with the Fresnel $F_0(\mathbf{p})$. In all our figures with SVBRDF parameter maps we therefore do not directly display $\mathbf{a}_s$, but in fact $\mathbf{a}_s(\mathbf{p}) F_0(\mathbf{p})$. For clarity, we still label it with $\mathbf{a}_s(\mathbf{p})$, though. In the main paper, we also drop the dependency on the surface point / pixel coordinates $\mathbf{p}$ for each of the parameters.

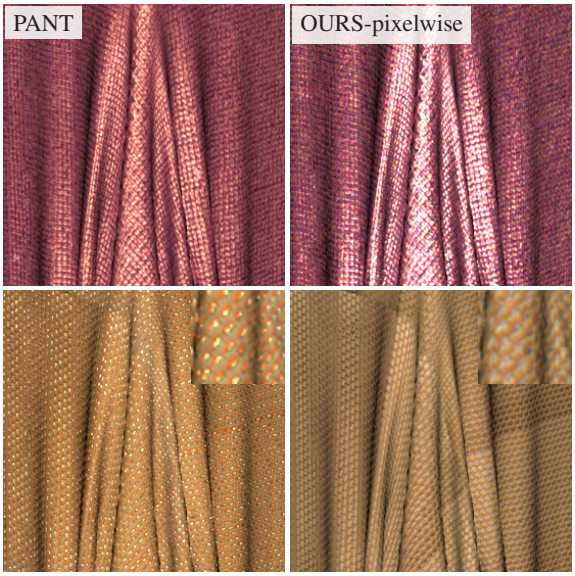

Figure 1: *Results obtained with our initial pixel-wise architecture. The lack of geometric context (neither neighborhood nor height) can cause an overestimation of highlights and an underestimation of albedos in shaded areas.*

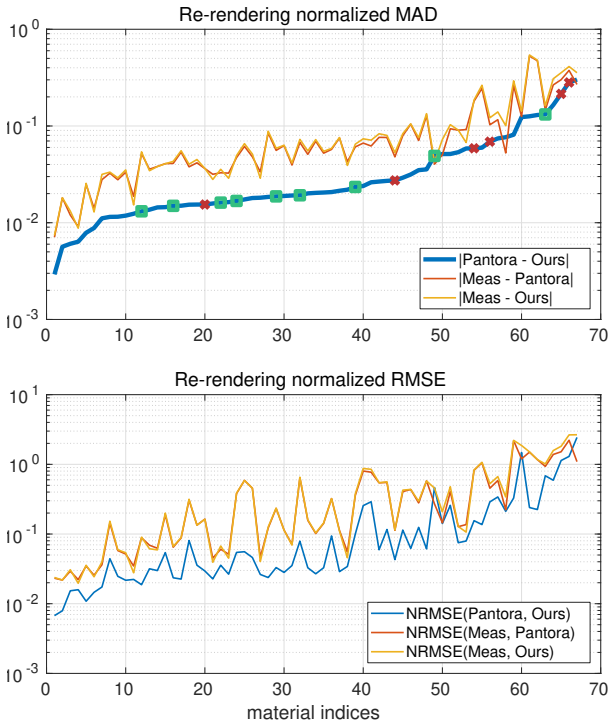

Figure 2: *Numeric results (normalized mean absolute deviation (MAD) and normalized RMSE) on all materials in our test set, sorted by increasing normalized MAD error between our estimates and the Pantora fits. Also plotted are the errors respectively between the measurements and the two fits. Highlighted with green boxes are the examples shown in Figures 5 & 6 in the main paper. The red crosses indicate the failure cases shown in Figures 8 & 9 in the main paper.*

4. Network Architecture

We performed our initial experiments with a pixel-wise network, set up as a multilayer perceptron (MLP) with a few fully connected layers. The inputs are, similar to Rainer et al. [RJGW19], the measured ABRDFs in each pixel. This MLP is fast to train and provides promising initial results. However, the predicted SVBRDFs consistenly showed artifacts in the form of overestimated glossiness and underestimated albedos in some areas, see upper row in Fig. 1. We explain these artifacts as a result of a lack of context over neighboring regions of the surface. If the current pixel has a lower height value than the neighboring areas, masking effects can make the observed heighlight appear sharper than it might be in an unoccluded scenario. Similarly, the observed diffuse component is underestimated due to shadowing. This can be seen in Fig. 1. To overcome these problems, we changed the inputs to patches, which provide the necessary neighborhood information.

5. Re-initialization Experiment

We conducted an experiment where we used our predicted SVBRDF parameters as initialization to the very last iteration of the Pantora SVBRDF fitting stage. For this, we limit the optimization iterations to 1 and feed and AxF file generated from our network predictions to the executable. This initializes the final surface geometry and the (achromatic) SVBRDF refinement stage. The subsequent per-pixel color fit is performed independent of the initialization. We observe convergence to virtually the same parameters. This final fitting stage takes roughly half the original fitting time, so any potential benefits are outweighed by a significant time penalty.

6. Additional Evaluation Results

Here we show additional numeric errors computed on the materials in our test split, similar to Fig. 4 in the main paper. Different to the main paper figure, we show both a *normalized* L_1 error (normalized mean absolute deviation, MAD), as well as a normalized L_2 error (normalized RMSE), to give a better impression of the distribution of errors. As normalization factors we use the maximum RGB intensity of the diffuse albedo of each material, as it provides a reliable normalization, insensitive to intensity variations caused by highlights. The results can be seen in Fig. 2. The figure also highlights the examples presented in Figures 5 – 8 in the main paper. Fig. 3 provides an overview over all materials in our test set, with similar annotations.

References

[GD10] GEISLER-MORODER, DAVID and DÜR, ARNE. "A new ward brdf model with bounded albedo". *Computer Graphics Forum*. Vol. 29. 4. Wiley Online Library. 2010, 1391–1398 1.

[GMMM14] GARRIDO-JURADO, SERGIO, MUÑOZ-SALINAS, RAFAEL, MADRID-CUEVAS, FRANCISCO JOSÉ, and MARÍN-JIMÉNEZ, MANUEL JESÚS. "Automatic generation and detection of highly

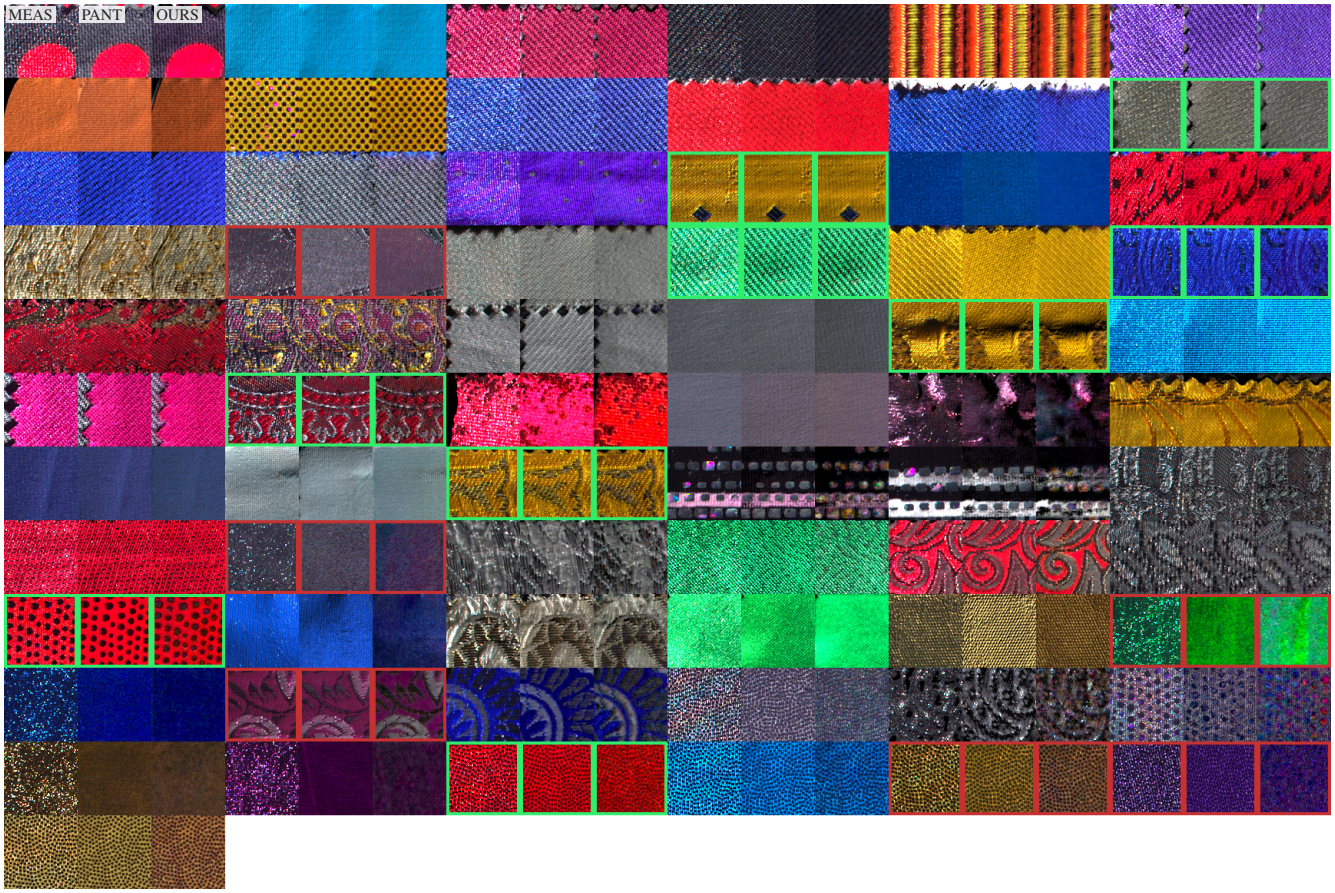

Figure 3: *Overview of our 67 test set materials, sorted in the same order as in Fig. 2. Highlighted by green frames are the results in Figures 5 & 6 in the main paper, by red frames the failure cases in Figures 8 & 9 in the main paper. The materials have the same ordering as in Fig. 2, each triple of cropped images shows in the following order: measured image (MEAS), re-rendered Pantora fit (PANT), re-rendering of our estimate (OURS).*

reliable fiducial markers under occlusion". *Pattern Recognition* 47.6 (2014), 2280–2292 1.

[RJGW19] RAINER, GILLES, JAKOB, WENZEL, GHOSH, ABHIJEET, and WEYRICH, TIM. "Neural BTF Compression and Interpolation". *Computer Graphics Forum (Proceedings of Eurographics)* 38.2 (Mar. 2019) 2.

[Sch94] SCHLICK, CHRISTOPHE. "An inexpensive BRDF model for physically-based rendering". *Computer graphics forum*. Vol. 13. 3. Wiley Online Library. 1994, 233–246 1.

[War*92] WARD, GREGORY J et al. "Measuring and modeling anisotropic reflection". *Computer Graphics* 26.2 (1992), 265–272 1.