

Gaussian Material Synthesis

Supplementary material

KÁROLY ZSOLNAI-FEHÉR, TU Wien

PETER WONKA, KAUST

MICHAEL WIMMER, TU Wien

ACM Reference Format:

Károly Zsolnai-Fehér, Peter Wonka, and Michael Wimmer. 2018. Gaussian Material Synthesis: Supplementary material. *ACM Trans. Graph.* 37, 4, Article 76 (August 2018), 5 pages. <https://doi.org/10.1145/3197517.3201307>

1 SHADER DESCRIPTION

Our shader is defined as a combination of a set of base BSDFs and mix shaders that return a linear interpolation of two inputs. In the following description, the numerical ID, BSDF model names and their parameters are enumerated.

1: *Diffuse BSDF* (r,g,b albedos), **2:** *Beckmann Glossy BSDF* (r,g,b albedos, roughness), **3:** *Mix shader* (connect 1,2), **4:** *Beckmann Glass BSDF* (r,g,b albedos, roughness, IOR), **5:** *Translucent BSDF* (r,g,b albedos), **6:** *Mix shader* (connect 4,5), **7:** *Mix shader* (connect 3,6).

The volume absorption for the Glass and Translucent BSDFs are inherited (and shared) from a separate node. The extended shader contains a combination of several noise models, a similar mixing logic and individual weighting factors to control. In the interest of simplicity, we provide a visual description of this shader in Fig. 2.

2 QUESTIONS & ANSWERS

Q: The new recommended samples are generated from scratch, and hence, have to be rendered. There is a large database of shader-image pairs that was used to train the neural network. Instead of synthesizing the corresponding images for new recommended sample points, why not use the images from the database?

A: We have experimented with this option and have included the appropriate source code and database file in our implementation. The issue of this approach is that it is struck with the curse of dimensionality; if one generates n training points on $[0, 1]^m$ with uniform distribution, and the average \mathcal{L}_2 distance between a new random query point and the closest database match is 0.71 when $m = 19, n = 17 \cdot 10^4$ (simpler shader) and 1.45 for $m = 38, n = 17 \cdot 10^4$ (extended shader). This difference is significant as some shader parameters are non-linear and a difference of 10^{-1} in the index of

Authors' addresses: Károly Zsolnai-Fehér, TU Wien, Favoritenstrasse 9-11/193-02, Vienna, Austria, 1040, zsolnai@cg.tuwien.ac.at; Peter Wonka, KAUST, Al Khwarizmi Bldg 1, Thuwal, Kingdom of Saudi Arabia, 23955-6900, pwonka@gmail.com; Michael Wimmer, TU Wien, Favoritenstrasse 9-11/193-02, Vienna, Austria, 1040, wimmer@cg.tuwien.ac.at.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2018 Copyright held by the owner/author(s).

0730-0301/2018/8-ART76

<https://doi.org/10.1145/3197517.3201307>

refraction or glass roughness leads to a drastically different material appearance or non-physical results. We note that we obtained these results via a Monte Carlo simulation as a complete proof of this result would be overly verbose (e.g., the average distance is close to 0.52 for $m = 2, n = 1 [1, 2]$).

Q: Why use 1D convolutions?

A: The input of the network is a set of shader parameters that bear some locality as e.g., R,G,B albedos for a material node are often adjacent. However, they have no meaningful 2D spatial relation to each other, therefore 1D convolutions are preferable.

Q: What about human biases in the gallery scores?

A: We have identified several recurring biases throughout our experiments. For instance, when the gallery is presented as a 2D grid, the score of a material often depends on its surroundings, e.g., it is typically rated higher when the material it surrounded by unfavorable examples. Users are also typically more susceptible to assign a high score to a mediocre sample early in the process before they have seen the best matches the shader has to offer. The perception of different light simulation effects may also introduce imperfections in the scoring process, e.g., the effect of translucency is difficult to recognize in moderation, and therefore there are barely any mid-scoring samples with most scores being either extremely low or high. In the Results section of the paper, we show that due to these biases, the Jensen-Shannon Divergence does not recede to zero as we add more training samples, and that despite these distortions, it is still possible to perform regression and material recommendations of formidable quality.

Q: What materials are synthesized and which are given in the right side of the teaser image with the Microplanet scene?

A: The following materials were synthesized: dandelions (upper part of the planet, high color variation), daisies (the white color is fixed, the core follows a slight color variation), staghorn tree (upper left), sweet pepper bush (lower right), Kentucky blue grass and rye (general vegetation covering the planet), the water steam in the middle (one material, extended shader). The following materials were given: the bark of the staghorn tree in the upper left, the procedural dirt material on the surface of the planet and the background HDR image.

Q: The neural network predictions are compared to the ground truth renderings in the paper. How do they compare against the training set images in terms of PSNR?

A: The PSNR values are comparable, i.e., 37.9dB for the predictions and 40.2dB for the training set when compared against the ground truth. In the paper, we argue that the visual quality of the predictions are higher because of the denoising property of the neural network. The measured PSNR is slightly lowered by the fact that some predicted images are off by a few points of brightness value;



Fig. 1. The *Microplanet* scene from the teaser image, magnified.

these are distributed uniformly over the image and are hence imperceptible for the user.

Q: In Fig. 5 of the paper with the Toy Tea Set scene, some of the material sample images seem slightly darker than the ones inserted into the scene. Why?

A: The volumetric absorption parameter for translucent materials is heavily scale-dependent: if the scale of the scene is larger than the one shown in the material preview scene (e.g., the teapots are typically larger), its increased optical thicknesses will result in darker outputs. We have slightly adjusted the absorptions to avoid this effect.

Q: It is noted that if the new recommendations are not acceptable, the user may rank the newly proposed gallery or edit past rankings. What is the difference?

A: If the recommendations are in line with the user's artistic vision, but require fine-tuning, assigning scores to the newly created gallery and using them as training data for one more round is expected to improve the relevance of the recommendations (we used a two-round scheme for metals and minerals). Since these new samples are *concatenated* to the previous scores, it is advisable to first make sure that the scores from the first round do not contain many conflicting decisions. If the recommendations are not at all acceptable, it is advisable to revisit the initial rankings.

Q: There is an abundance of glassy and translucent materials in the provided examples. Why?

A: These cases are considered challenging in the sense that these

materials are relatively unlikely to appear via random sampling: in the glassy use case, 81% of the samples in the initial gallery were scored zero. This ratio was 90% for the translucent case. This means that the recommender system has to learn the appropriate sample distribution from a modest number of non-zero data points. With this, we intended to show that our system performs well in the more challenging cases. Other material classes that represent a larger slice of the parameter space (e.g., highly diffuse materials) are learned faithfully from even fewer training samples.

Q: How would testing more novice and expert users change the modeling timings?

A: We have found that the fixed cost of the direct interaction with a principled shader is consistent among users. Our expert and novice users noted that most of their time was spent waiting for noise to clear up in the rendered images when a parameter is changed. This effect is particularly pronounced during variant generation, where the user has to wait until the minute differences between the old and new variant are revealed. This is a shortcoming that is inherent in the fact that all images have to be re-rendered and remains true for all users.

Q: Are the shader-image pairs in the training data sampled with uniform distribution? Could this be improved?

A: Yes, we have used uniform sampling – the accuracy of the neural renderer can be further improved through an adaptive, non-uniform sampling of the parameter space for the training set. The fact that even the uniform case led to satisfactory results accentuates the

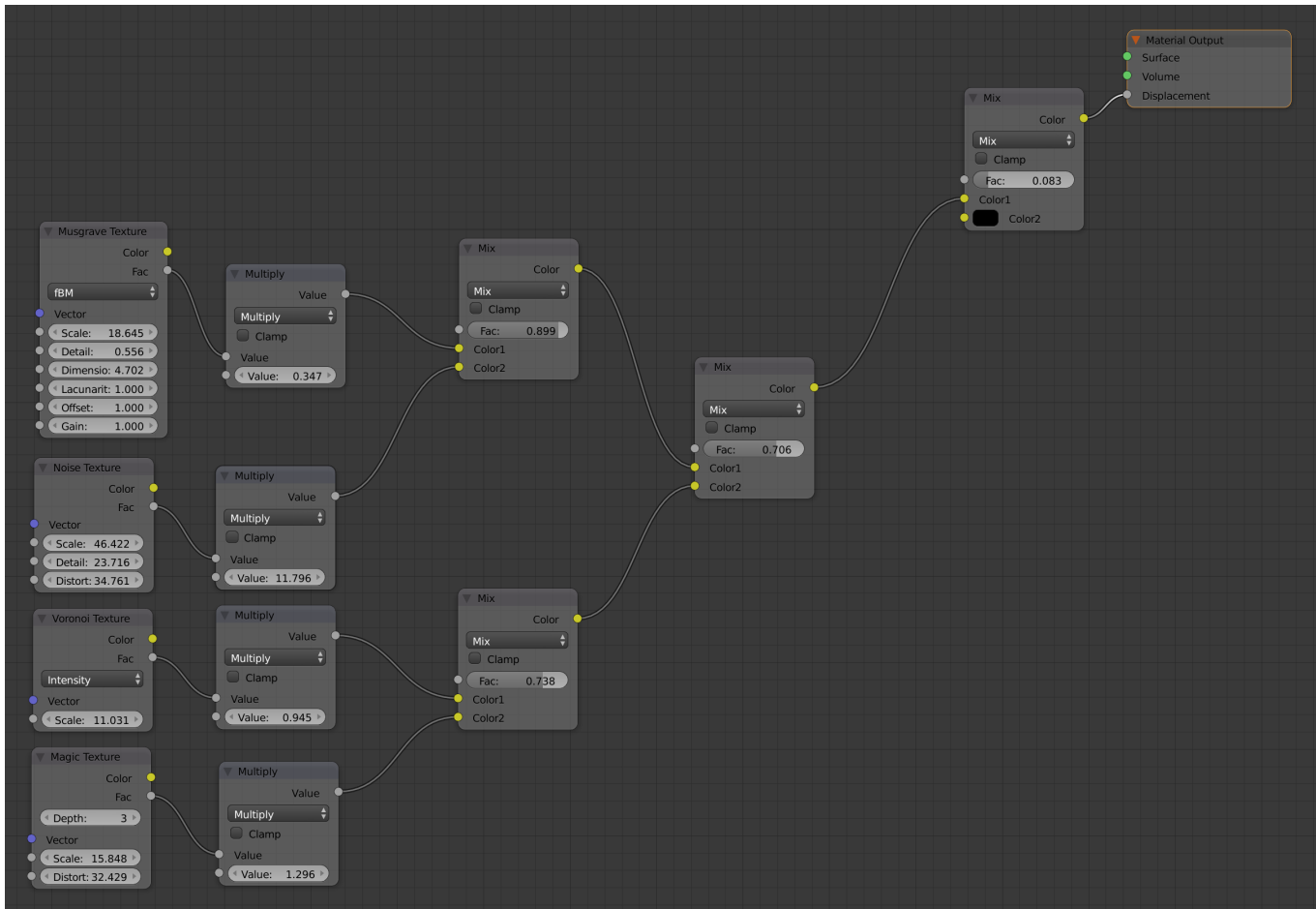


Fig. 2. Our principled shader for generating a wide variety of possible displacements.

utility of the neural rendering step. Furthermore, uniform sampling of new principled shaders can take place conveniently without additional complexity with few adaptations to our implementation.

Q: Would it be possible to explore the denoising property of the neural network in a more principled way?

A: The noise filtering property of the neural network is indeed subject to a tradeoff – adding more layers leads to more faithfully rendered images at the risk of additionally fitting the noise in the dataset. We think that a more principled approach could be developed by using modern neural network visualization techniques to observe the amount of noise contained within the filters [3].

3 SUPPLEMENTAL FILES

The submission contains a supplementary video with a high-level overview of our system and a discussion of the results. We have also attached a compressed archive that contains the following materials: GPR training data and full workflows for the glassy, translucent, metals and minerals and glittery materials accompanied by 500 rendered images in the gallery that can be scored by the user, and a Blender scene containing the description of our principled shader.

Additionally, the code for computing the answer to the first question in Section 1., and 1D active learning experiments from the Future Work section of the paper will also be included to foster further research in this area.

REFERENCES

- [1] Bernhard Burgstaller and Friedrich Pillichshammer. 2009. The average distance between two points. *Bulletin of the Australian Mathematical Society* 80, 3 (2009), 353.
- [2] AM Mathai, P Moschopoulos, and G Pederzoli. 1999. Random points associated with rectangles. *Rendiconti del Circolo Matematico di Palermo* 48, 1 (1999), 163–190.
- [3] Chris Olah, Arvind Satyanarayan, Ian Johnson, Shan Carter, Ludwig Schubert, Katherine Ye, and Alexander Mordvintsev. 2018. The Building Blocks of Interpretability. *Distill* (2018). <https://doi.org/10.23915/distill.00010> <https://distill.pub/2018/building-blocks>.



Fig. 3. The *Toy Tea Set* scene showcasing translucent material models learned by our technique.

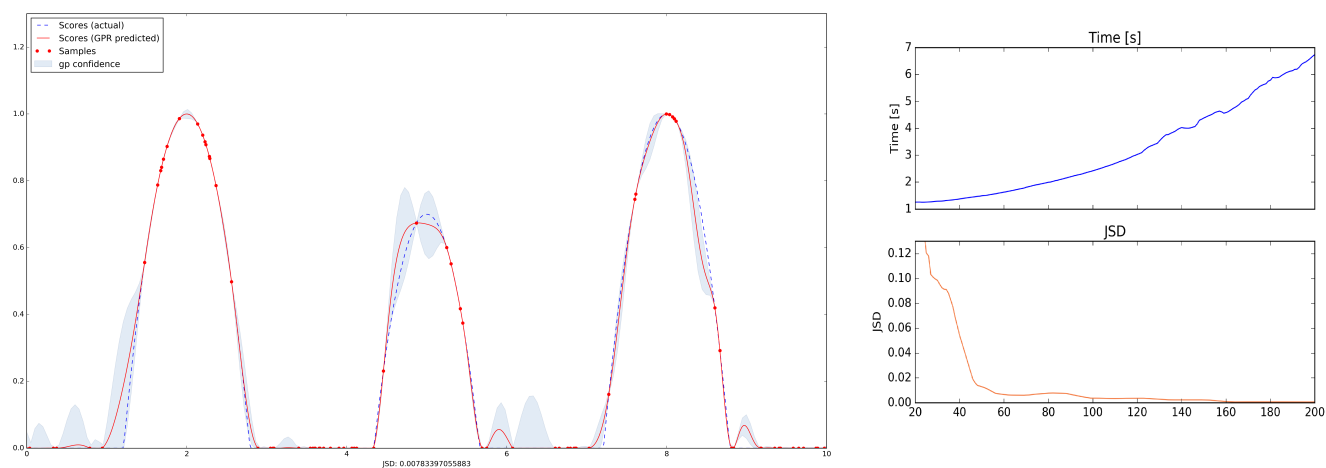


Fig. 4. Gaussian Process Regression in 1D and the corresponding JSD and execution timings.



Fig. 5. The *Still Life* scene from the teaser image, magnified.