Supplementary Materials for PhotoMat: A Material Generator Learned from Single Flash Photos

Xilong Zhou
Texas A&M University
Adobe Research
College Station, USA

Miloš Hašan
Adobe Research
San Jose, USA

Valentin Deschaintre
Adobe Research
London, UK

Paul Guerrero
Adobe Research
London, UK

Yannick Hold-Geoffroy
Adobe Research
San Jose, USA

Kalyan Sunkavalli
Adobe Research
San Jose, USA

Nima Khademi Kalantari
Texas A&M University
College Station, USA

CCS CONCEPTS

• **Computing methodologies → Reflectance modeling**.

KEYWORDS

Materials, SVBRDF, generative models, GAN

ACM Reference Format:
Xilong Zhou, Miloš Hašan, Valentin Deschaintre, Paul Guerrero, Yannick Hold-Geoffroy, Kalyan Sunkavalli, and Nima Khademi Kalantari. 2023. Supplementary Materials for PhotoMat: A Material Generator Learned from Single Flash Photos . In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Proceedings (SIGGRAPH '23 Conference Proceedings), August 6–10, 2023, Los Angeles, CA, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3588432.3591535

In these supplementary materials, we visually demonstrate the accuracy of light detection of our training dataset, additional visual comparison of PhotoMat and TileGen, more curated sampling results, non-curated sampling results and inverse rendering results.

Light detection. As described in the main paper, we design a simple approach to detect the flash light position for our training images. In Fig 1, we mark the detected light position as a red square in each example to illustrate the accuracy of light detection. As the figure shows, obvious highlights are correctly found, while non-obvious cases are resolved plausibly.

Comparison against TileGen. In Fig. 2, we demonstrate a side by side visual comparison of PhotoMat and TileGen. Since TileGen is trained per material category, we only compared stone and leather materials. For PhotoMat, we select stone and leather from both 256^2 and 512^2 model. For each category we add 30 examples for both PhotoMat and TileGen to the figure. As is seen, the materials generated from PhotoMat are visually more realistic compared with the materials generated from TileGen, which is consistent with the results of user study.

More results. In Fig 3 and Fig 4, we demonstrate relit neural materials, analytic rendered materials and estimated SVBRDF of 256^2 model. Similarly, in Fig 5 and Fig 6 we demonstrate more results of 512^2 model. Finally, in Fig 7 we show results generated from our 1024^2 model.

Non-curated samples. In Fig 8 and Fig 9, we demonstrate non-curated (random) sampling results of our pretrained 256^2 and 512^2 model. More specifically, we show 120 independently randomly sampled materials for each model.

Inverse rendering. Similar to MaterialGAN [**?**] and TileGen [**?**], PhotoMat can be used to achieve material acquisition aiming at style similarity given a single photo. We use the tileable 256^2 model and follow the strategy proposed by TileGen to achieve inverse rendering. More specifically, we utilize the pretrained generator **G** and maps estimator **E** as priors, and optimize latent and noise space under Gram matrix loss with shift strategy, following TileGen. As shown in Fig. 10, the resulting re-rendered images match the style of the targets and the estimated SVBRDFs do not suffer any highlight burn-in artifacts, demonstrating the applicability of PhotoMat to inverse rendering.

SIGGRAPH '23 Conference Proceedings, August 6–10, 2023, Los Angeles, CA, USA

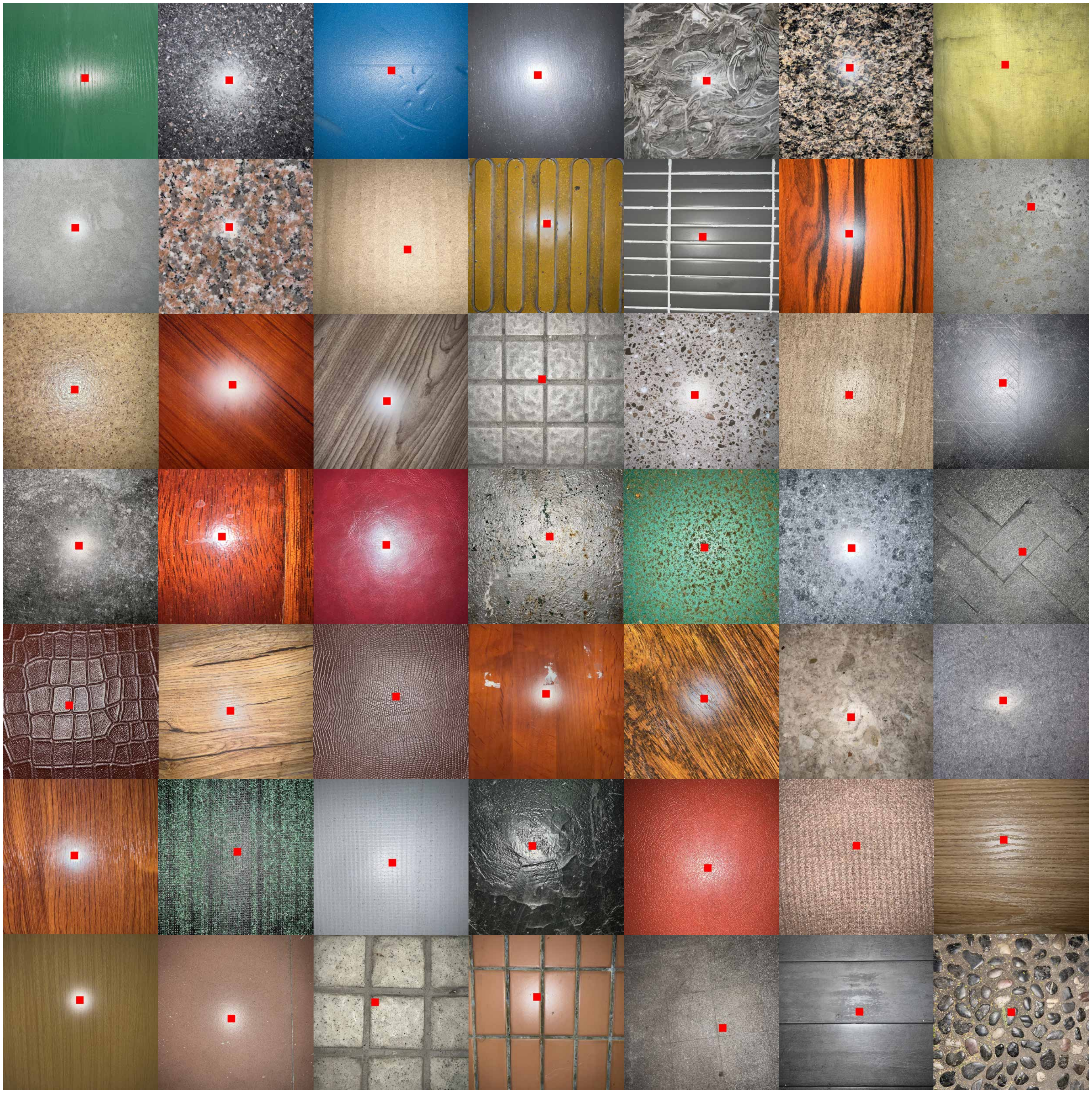

Figure 1: We illustrate the accuracy of detected light position for our training dataset. The detected light is marked as pure red square in each example and this matches the real flash light position, which demonstrate the accuracy of our light detection method.

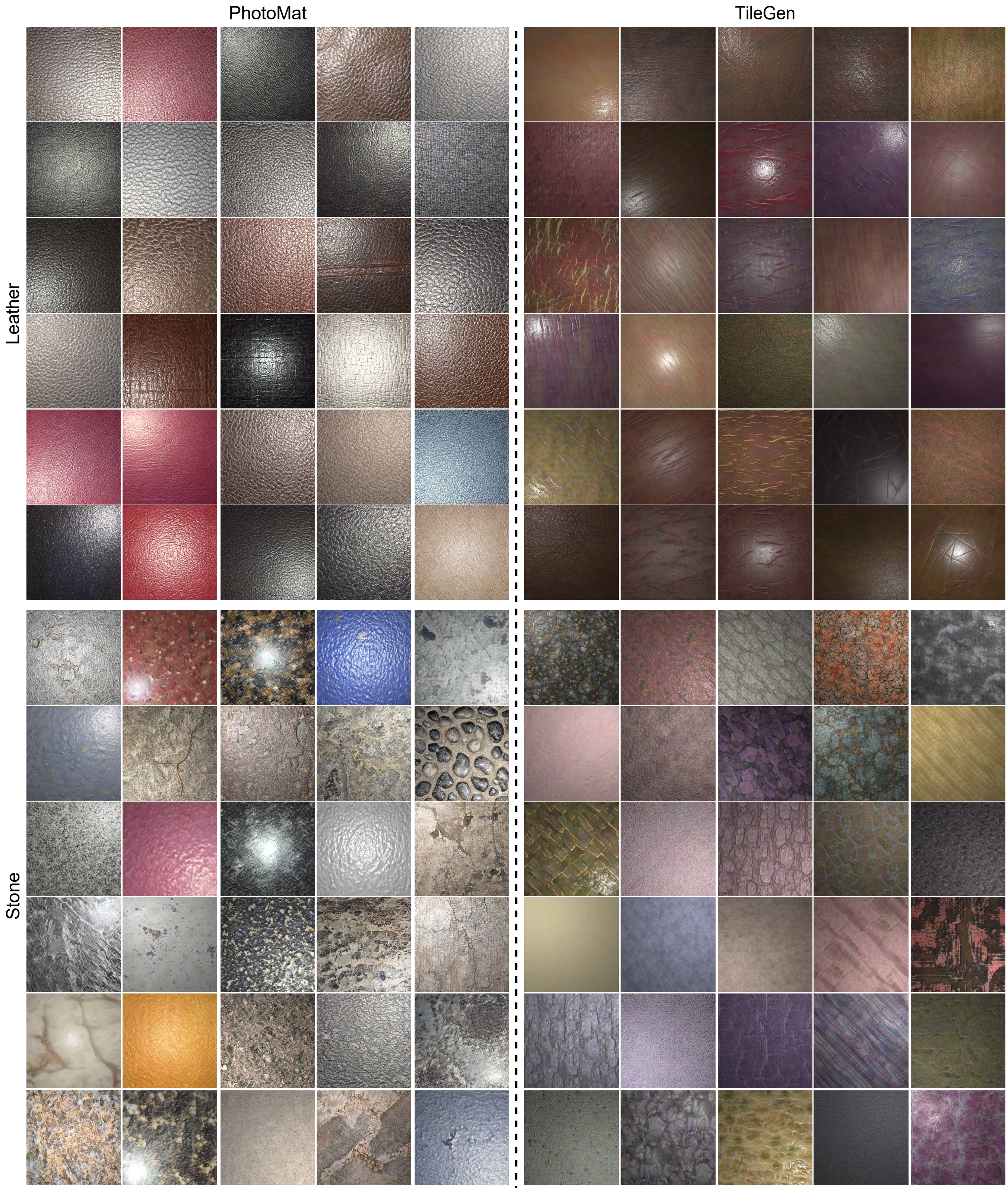

Figure 2: We demonstrate a visual comparison of PhotoMat (left) and TileGen (right) on leather (top) and stone (bottom) materials.

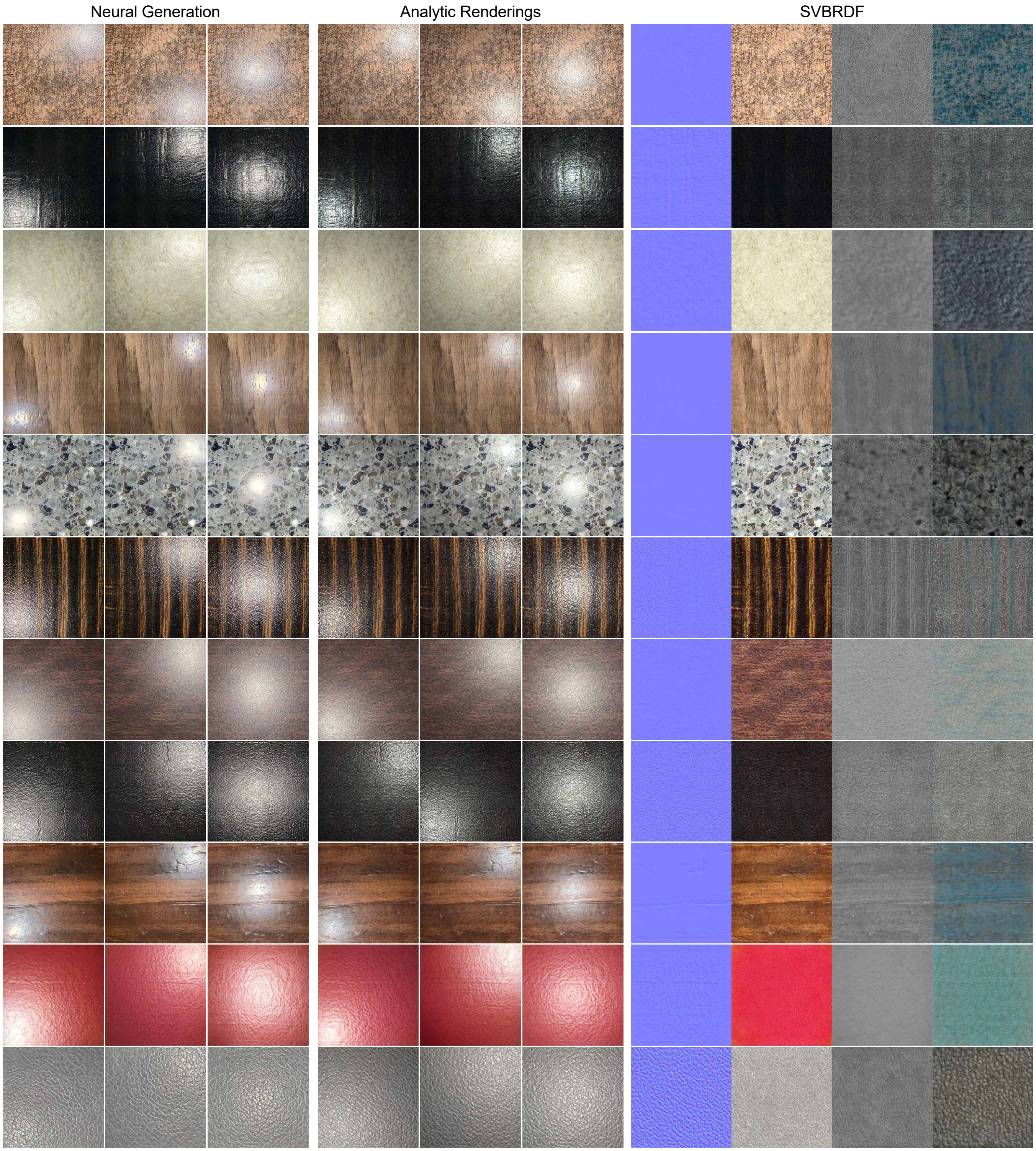

Figure 3: More results generated from 256^2 model. From left to right: relit neural materials, analytic rendered materials and generated SVBRDF.

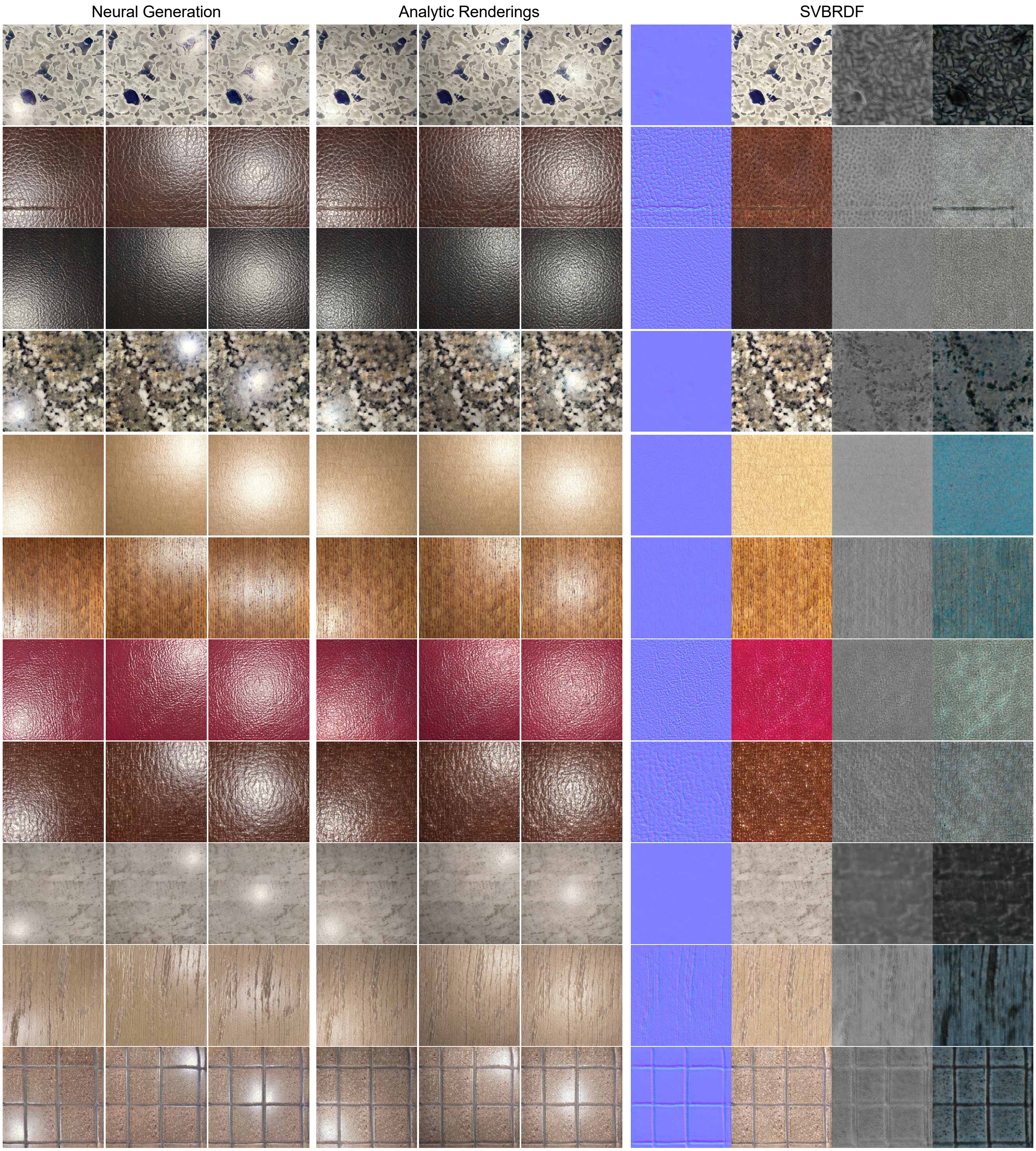

Figure 4: More results generated from 256^2 model. From left to right: relit neural materials, analytic rendered materials and generated SVBRDF.

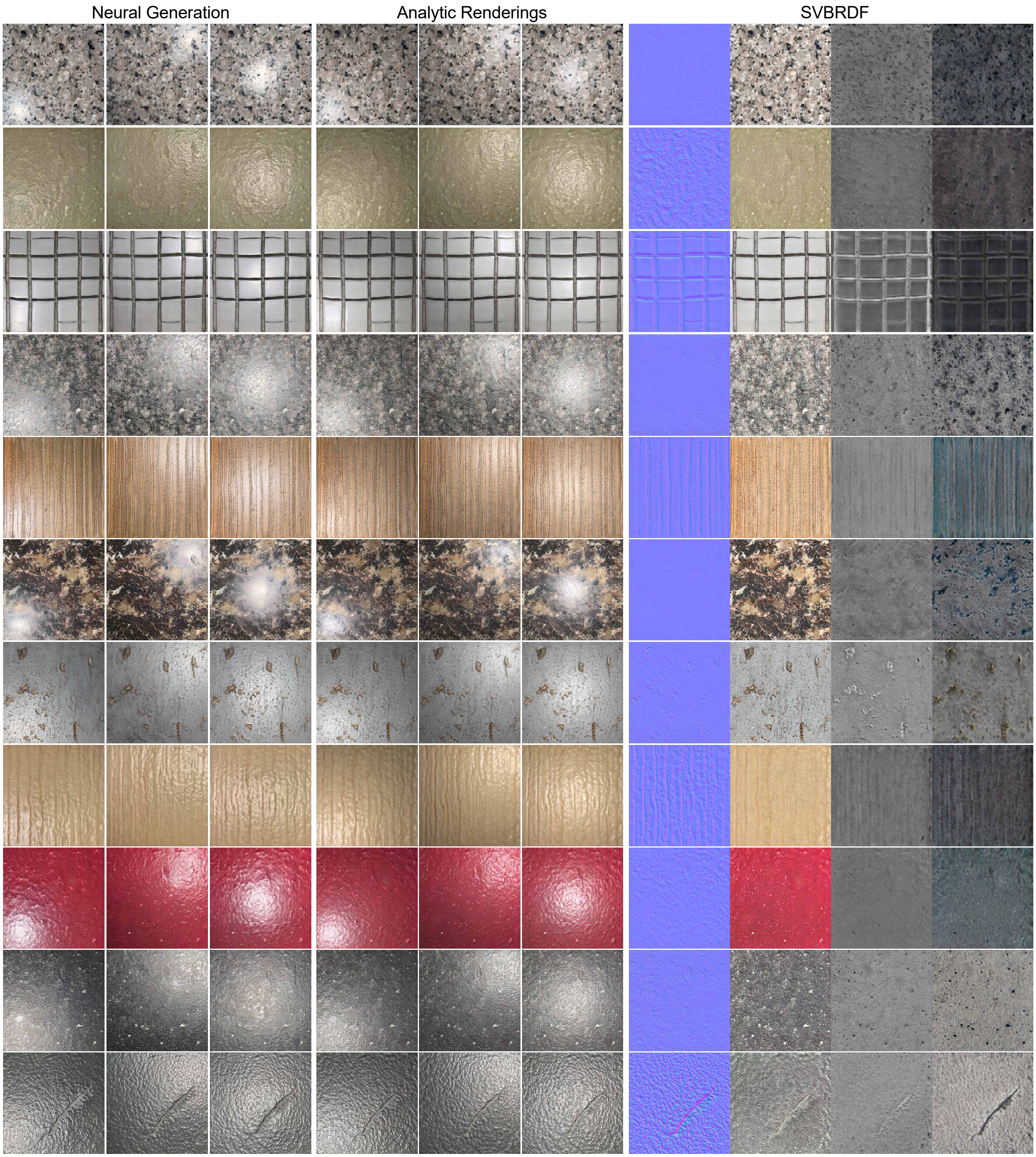

Figure 5: More results generated from 512^2 model. From left to right: relit neural materials, analytic rendered materials and generated SVBRDF.

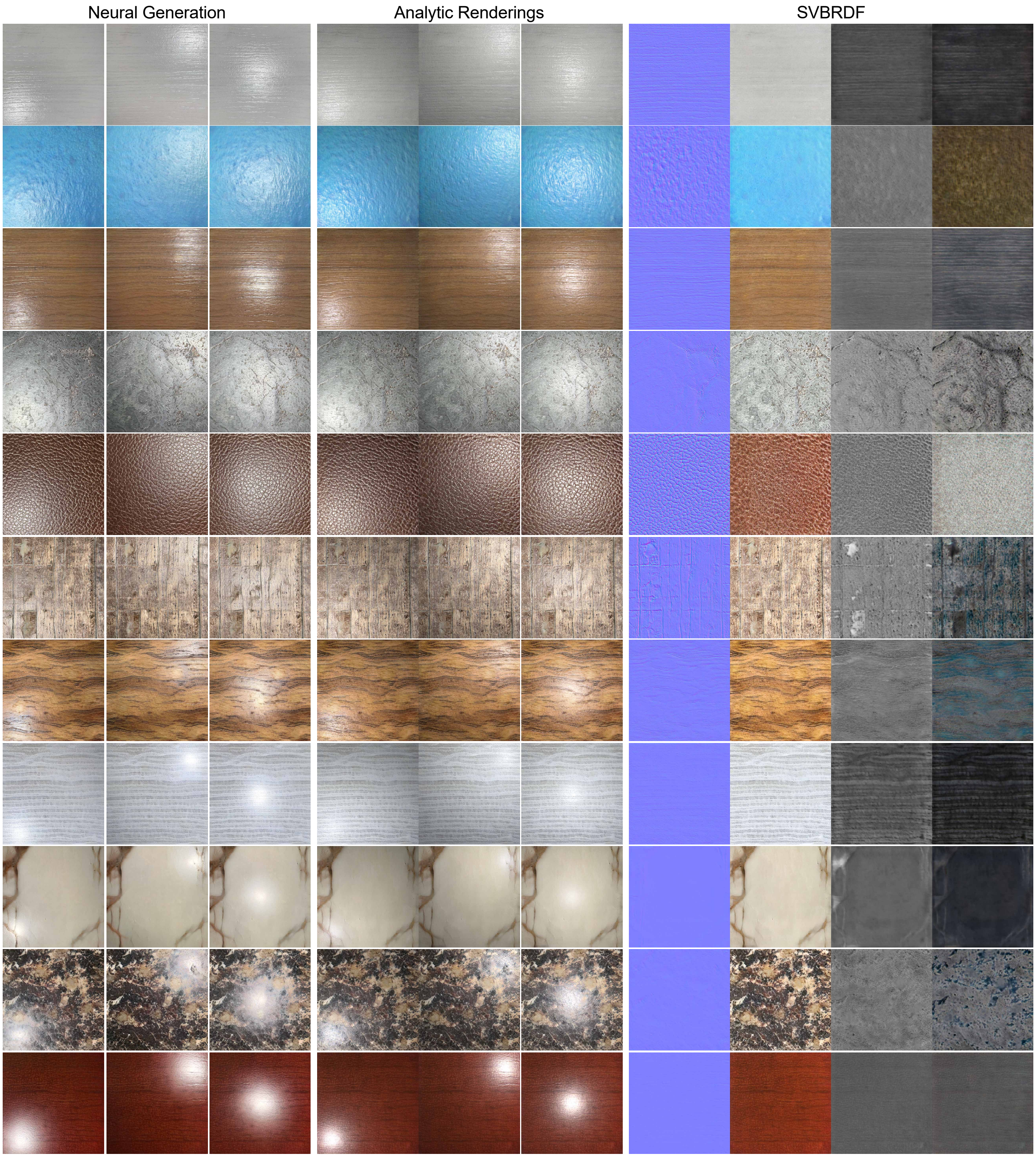

Figure 6: More results generated from 512^2 model. From left to right: relit neural materials, analytic rendered materials and generated SVBRDF.

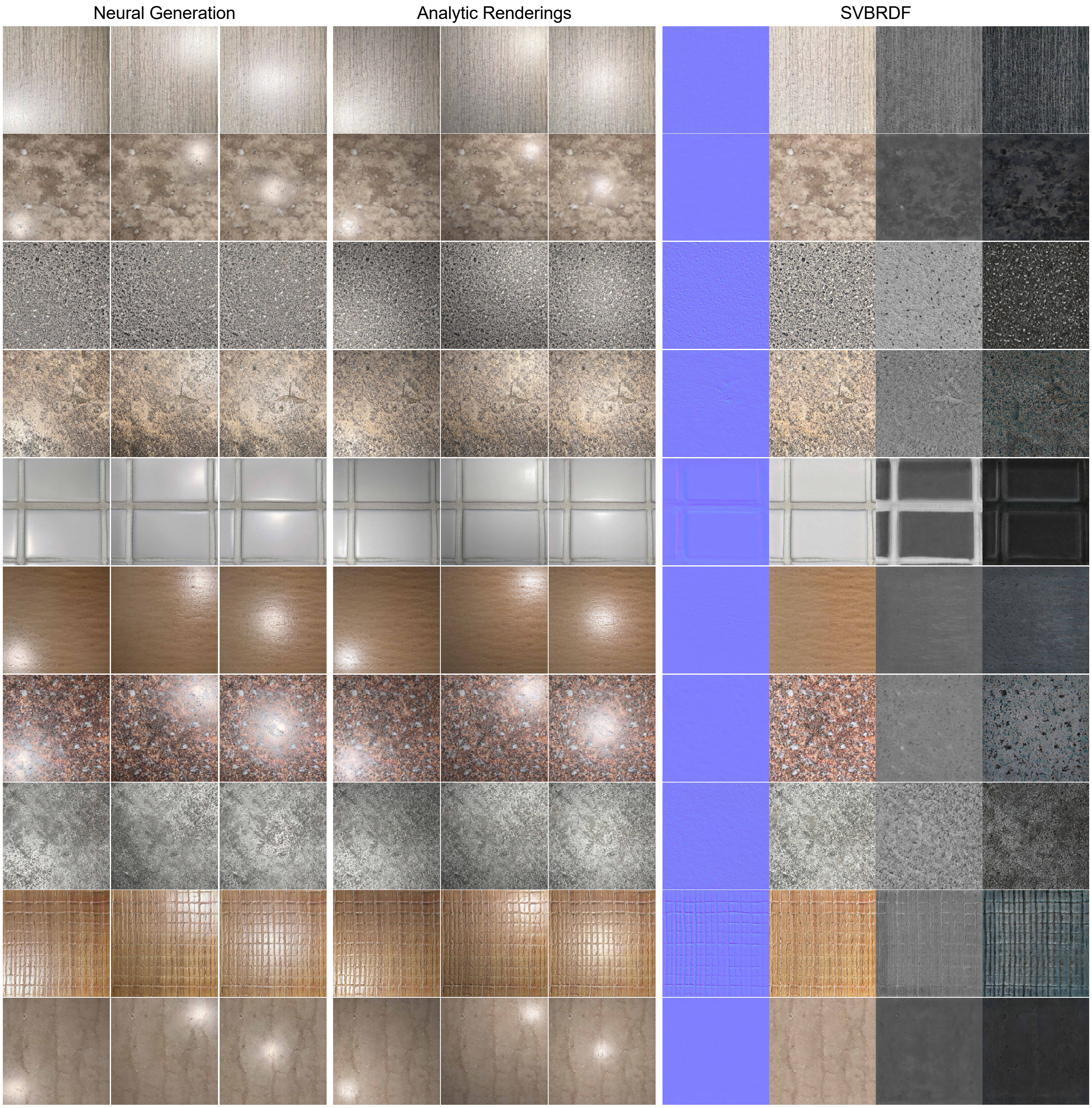

Figure 7: More results generated from 1024^2 model. From left to right: relit neural materials, analytic rendered materials and estimated SVBRDF.

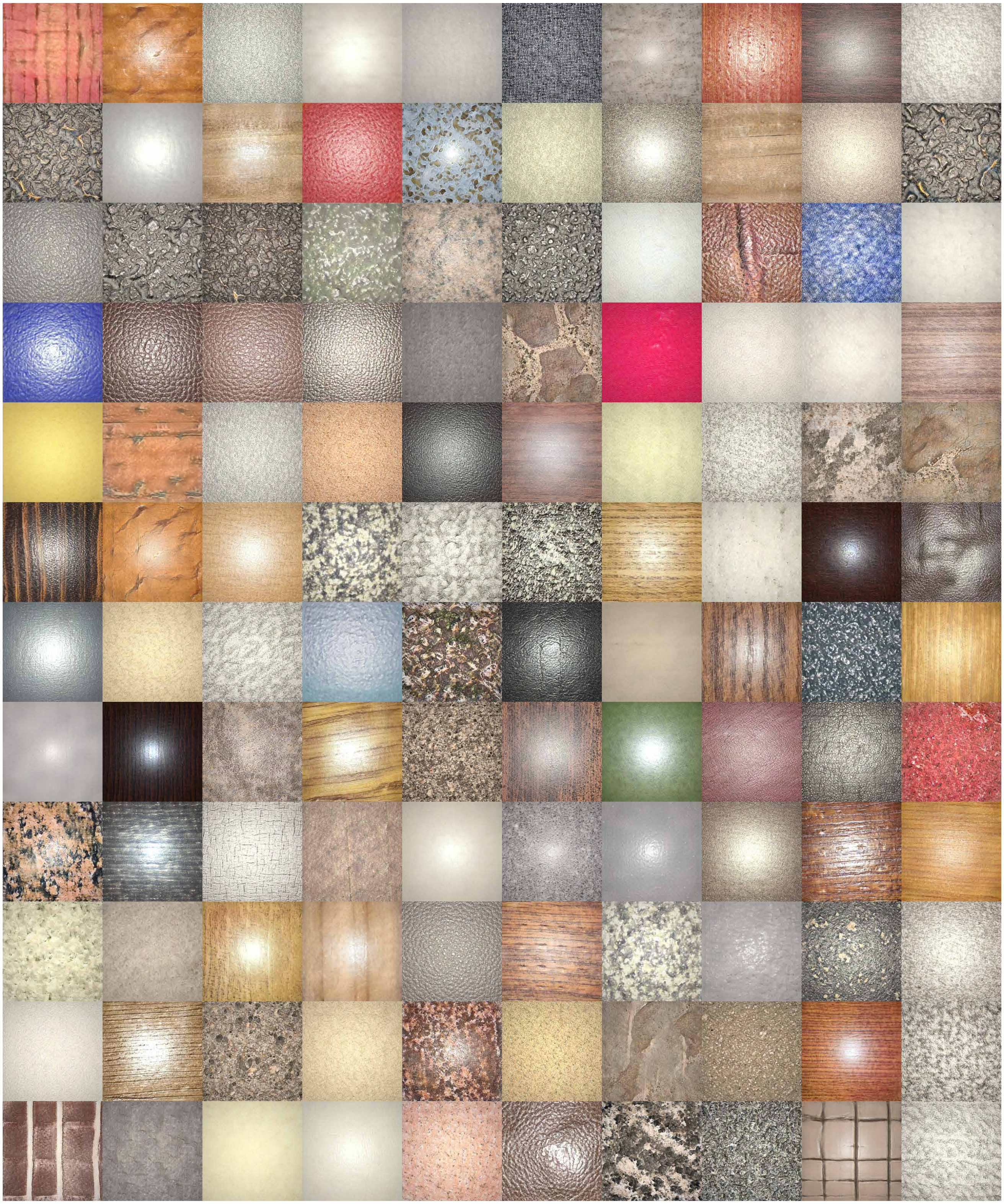

Figure 8: Non-curated 120 samples from 256^2 model trained on the small *Glossy* dataset.

Figure 9: Non-curated 120 samples from 512^2 model trained on the large "in-the-wild" dataset.

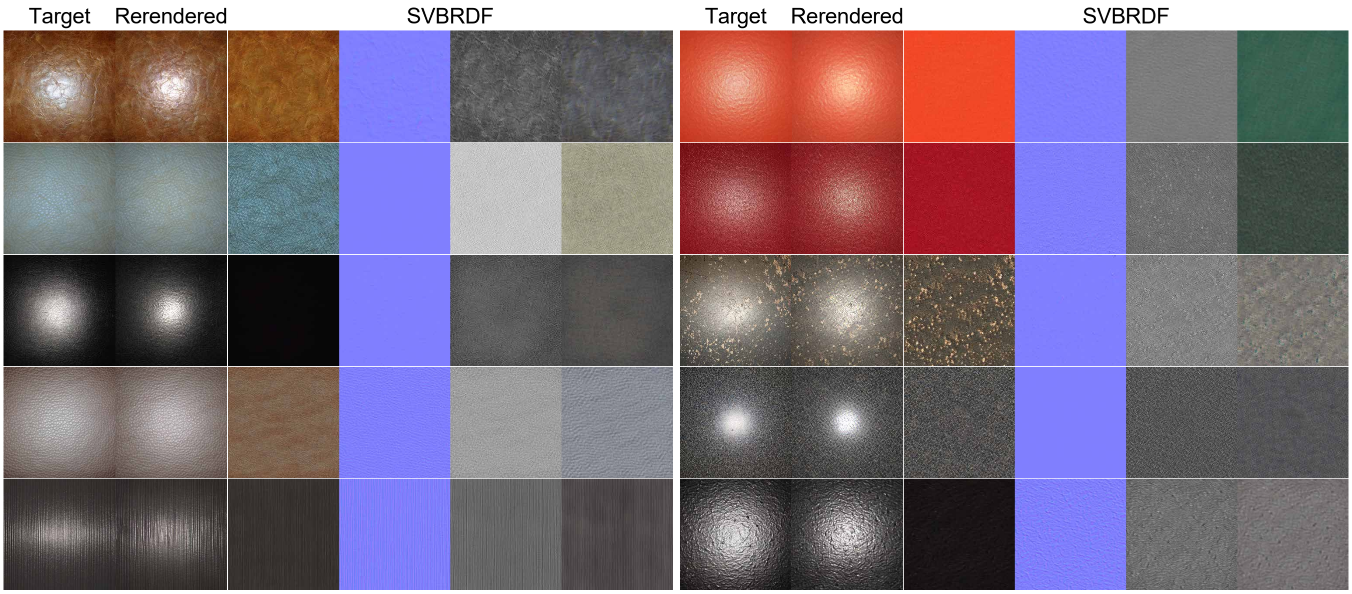

Figure 10: Inverse rendering results by optimizing the latent and noise space to match the target photos under Gram matrix loss. As shown here, PhotoMat can achieve high quality SVBRDF maps without burn-in artifacts.