M3ashy: Multi-Modal Material Synthesis via Hyperdiffusion

Chenliang Zhou, Zheyuan Hu, Alejandro Sztrajman, Yancheng Cai, Yaru Liu, Cengiz Öztireli.

Department of Computer Science and Technology, University of Cambridge.

https://peterhuistyping.github.io/M3ashy/

Contributions

Two new material datasets,

multi-modal synthesis of data-driven material appearance models,

a novel statistics-based constrained synthesis,

two BRDF distributional evaluation metrics,

further ablation studies are conducted.

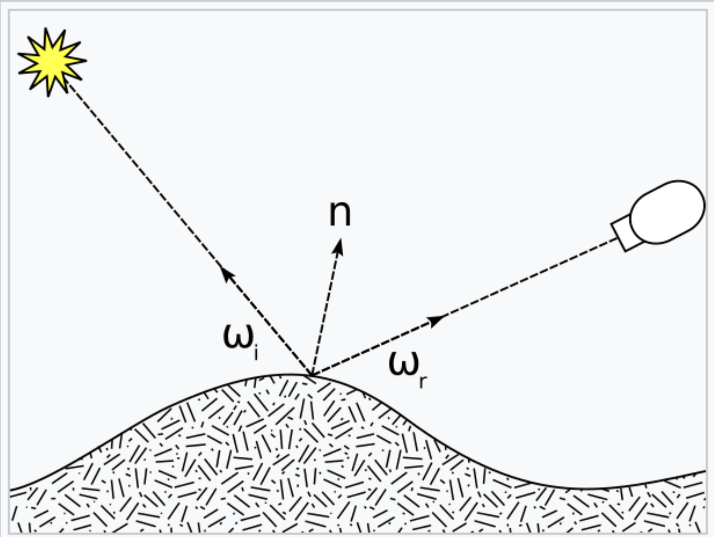


Diagram
BRDF.
toward
viewer

https://en.w

Introduction and Motivation

The rendering equations

$$L_{\mathrm{o}}(\mathbf{x},\omega_{\mathrm{o}},\lambda,t) = L_{\mathrm{e}}(\mathbf{x},\omega_{\mathrm{o}},\lambda,t) + L_{\mathrm{r}}(\mathbf{x},\omega_{\mathrm{o}},\lambda,t)$$

$$L_{\mathrm{r}}(\mathbf{x},\omega_{\mathrm{o}},\lambda,t) = \int_{\Omega} f_{\mathrm{r}}(\mathbf{x},\omega_{\mathrm{i}},\omega_{\mathrm{o}},\lambda,t) L_{\mathrm{i}}(\mathbf{x},\omega_{\mathrm{i}},\lambda,t)(\omega_{\mathrm{i}}\cdot\mathbf{n})\,\mathrm{d}\,\omega_{\mathrm{i}}$$

BRDF modeling method	Measured	Generative	Type	Text	Image	CS	Datasets	Metrics
DeepBRDF (Hu et al. 2020)	✓	✗	✗	✗	✓	✗	✗	✗
Henzler et al. (2021)	✗	✓	✗	✗	✓	✗	✗	✓
MATLABER (Xu et al. 2023)	✗	✓	✗	✓	✗	✗	✗	✗
Memery, Cedron, and Subr (2023)	✗	✓	✗	✓	✗	✗	✗	✗
Gokbudak et al. (2023)	✓	✗	✗	✗	✓	✗	✗	✗
M^{3}ashy (ours)	✓	✓	✓	✓	✓	✓	✓	✓

Table 1: Comparison of material modeling methods. Our M^{3}ashy is the first generative pipeline for measured real-world materials that supports both unconditional and multi-modal conditional synthesis guided by type, text, or image. It also enables a statistics-based constrained synthesis (CS) and introduces novel datasets and material distributional metrics.

Dataset adopted: MERL

- MERL dataset [MPBM03], where each BRDF is stored using $90 \times 90 \times 180 \times 3$ floating point numbers.
    - diverse, a total of 100, real-world captured materials.
    - data-driven property.
        - suitable for applying statistical or neural network-based methods on them.
- assume the materials are isotropic and spatially-identical for simplicity.
    - possible extensions in the future work: considering spatially-varying materials.

https://cdfg.csail.mit.edu/wojciech/brdfdatabase

Material synthesis **framework**

1 (top left): ***Aug**MERL* dataset.

- Data augmentation using RGB permutation and PCA interpolation to create an expanded dataset.

2 (middle): ***Neu**MERL* dataset.

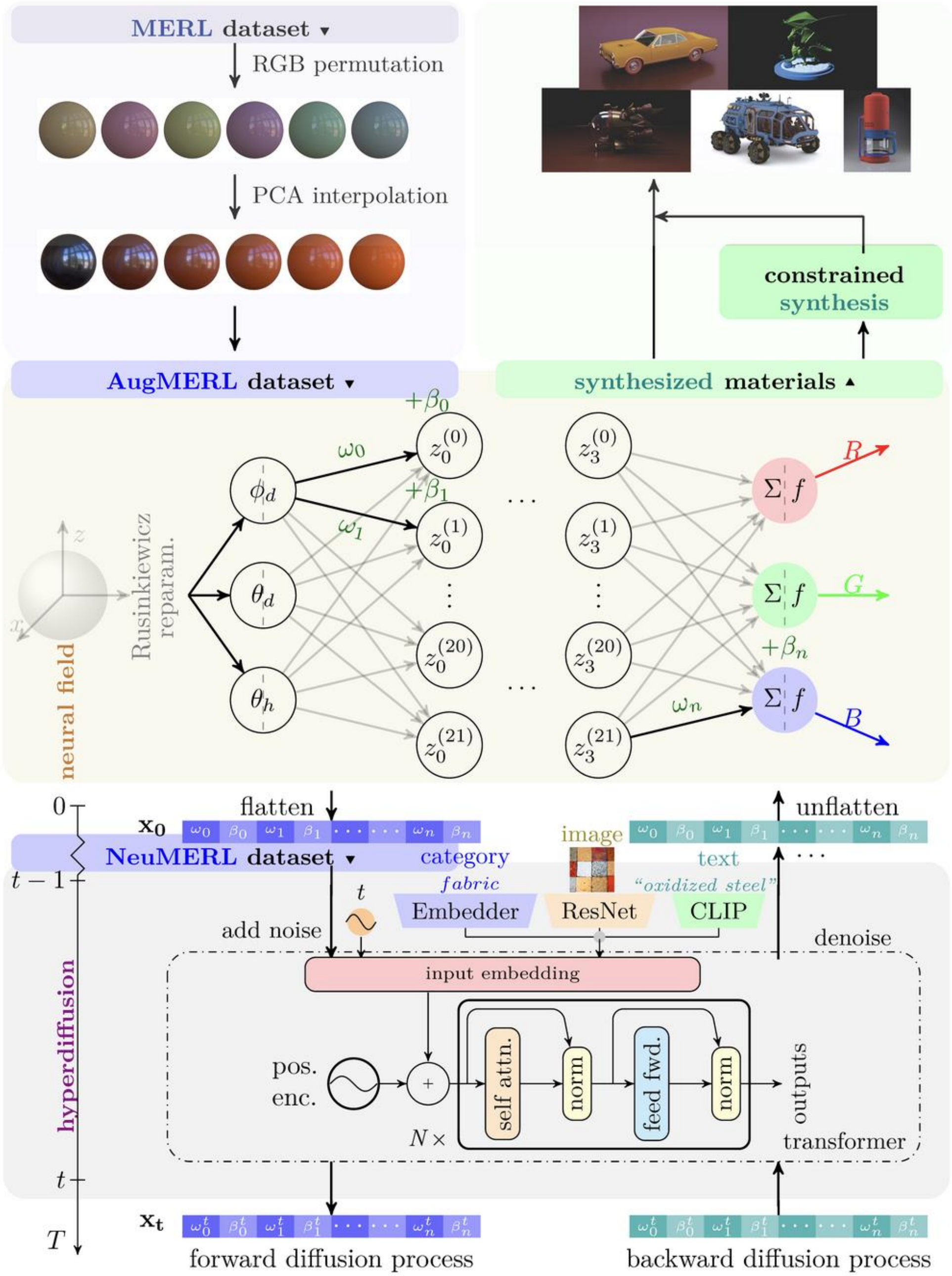
- Neural field fitted to individual materials, resulting in *NeuMERL*, a dataset of neural material representations;

3 (bottom): ***Hyperdiffusion*** on NeuMERL.

- Training a multi-modal conditional hyperdiffusion model on NeuMERL to enable conditional synthesis of materials guided by inputs such as material type, text descriptions, or reference images.

4 (top right) ***Statistics***-based constrained synthesis.

- to generate materials of a specified type.



Unconditional synthesis

- Evaluation with different baselines.
- Demonstration of complex visual Graphics results.

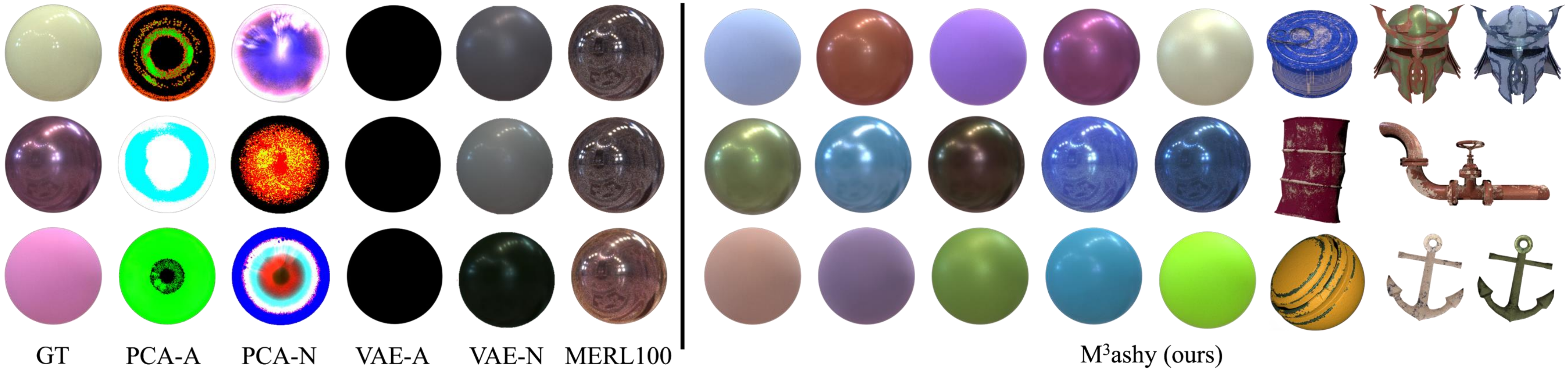
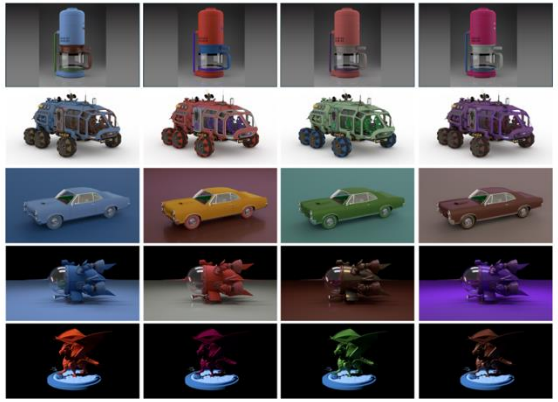


Figure 20: Our synthesized neural materials rendered with bump maps.

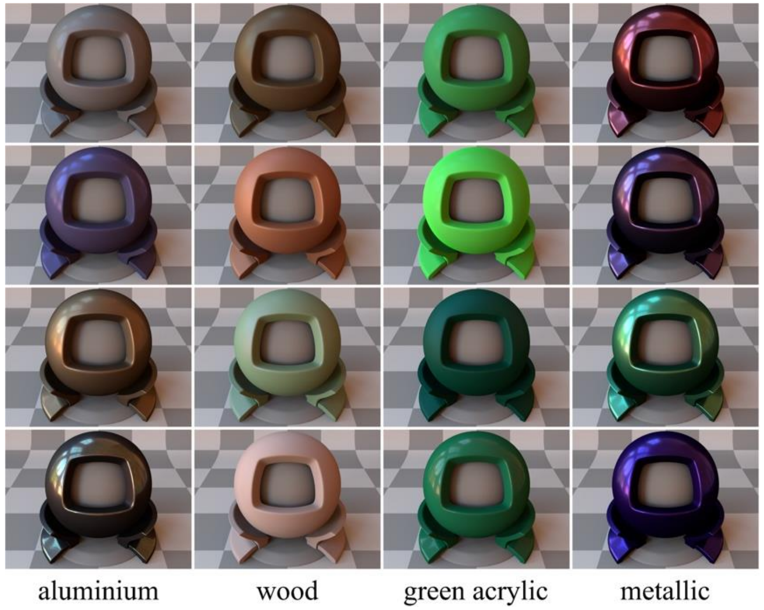


Additional rendering results

- Renderings of different 3D models using our synthesized neural materials, highlighting the quality and diversity.

Novel metrics for material synthesis

Metric		Training set	PCA-A	PCA-N	VAE-A	VAE-N	MERL100	M^3ashy (ours)
FID (↓)		0.187	10.9	23.8	26.1	10.0	7.56	**0.440**
MMD (↓)	BRDF-L1$\times 10^{-3}$	2.51	9.05	9.22	9.09	5.83	4.30	**4.02**
	RMSE$\times 10^{2}$	7.54	33.3	30.2	63.7	15.5	13.4	**9.34**
	NegPSNR	−28.7	−13.9	−14.8	−8.30	−20.9	−22.6	**−25.6**
	NegSSIM$\times$10	−9.55	−6.74	−6.29	−2.68	−6.86	−8.27	**−9.40**
COV (%) (↑)	BRDF-L1	60.8	2.50	30	0.833	20.8	28.3	**50.8**
	RMSE	55.8	18.3	28.3	0.833	16.7	25.0	**50.0**
	NegPSNR	56.7	18.3	28.3	0.833	18.3	25.0	**50.0**
	NegSSIM	59.2	23.3	16.7	0.833	17.5	22.5	**51.7**
1-NNA (%) (↓)	BRDF-L1	58.8	100	95.4	100	96.7	92.5	**80.0**
	RMSE	55.4	96.3	93.4	100	93.3	84.6	**60.0**
	NegPSNR	55.0	94.2	90.0	100	93.3	84.6	**60.4**
	NegSSIM	57.5	96.3	96.7	100	93.8	86.3	**61.7**



Multi-modal synthesis evaluation

- Left: **Text**-conditioned,
  - including unseen text labels, e.g. "green metal", "red plastic", and "highly specular material".
- Middle: **Type**-conditioned.
- Right: **Image**-conditioned.
  - reference condition | synthetic material.

Constrained synthesis evaluation

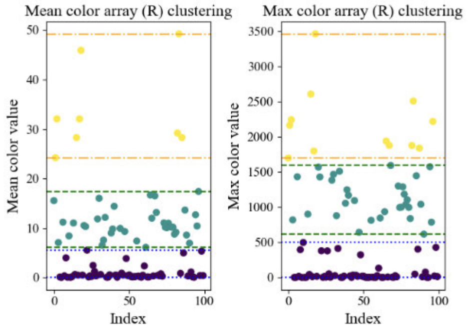
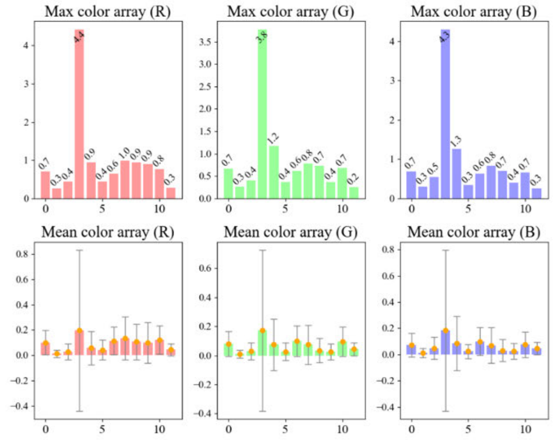
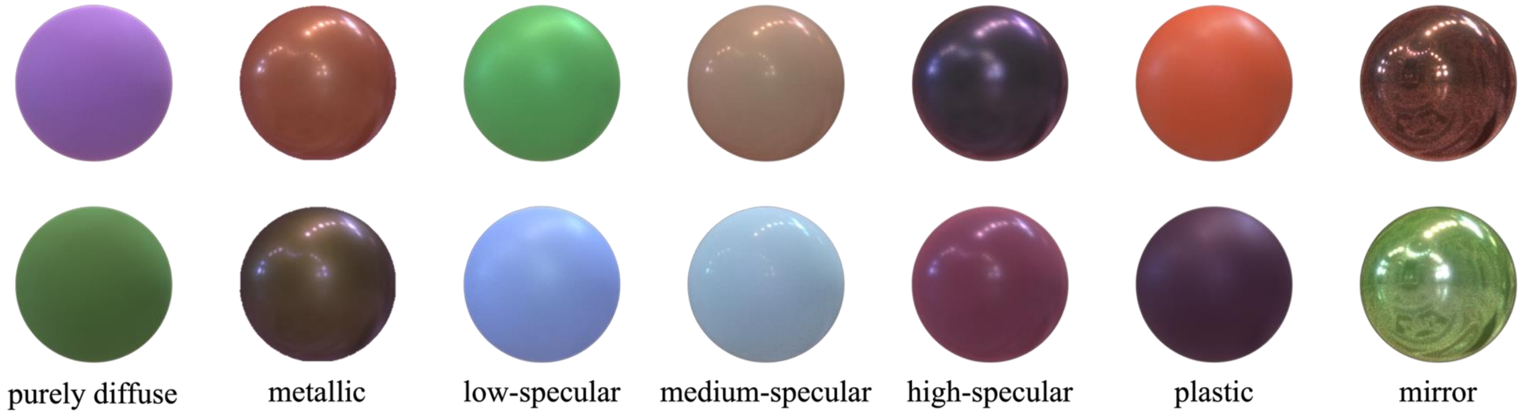
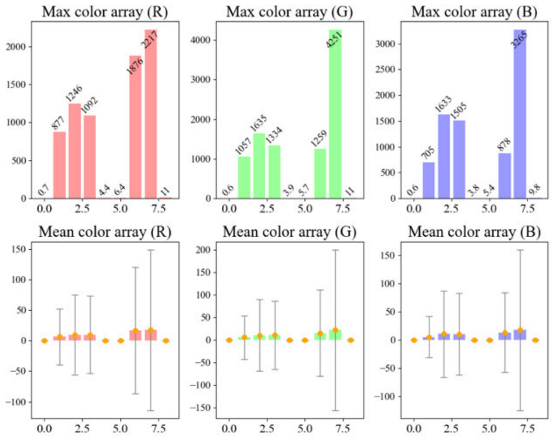


Figure 25: K-means clustering on mean (left) and maximum (right) reflectance of red channel.



(a) Fabric BRDF statistics





(c) Plastic BRDF statistics

Ablation Study

Metric	Sparse reconstruction		Compression	
	MERL	AugMERL	MERL	AugMERL
PSNR (↑)	32.2	**36.3**	45.2	**48.3**
Delta E (↓)	2.1	**1.8**	0.693	**0.623**
SSIM (↑)	0.972	**0.983**	**0.994**	**0.994**

Table 3: Quantitative comparison of training on MERL versus AugMERL in the sparse BRDF reconstruction and BRDF compression experiments. The results demonstrate that training on AugMERL consistently enhances performance across all metrics.

Reference: *Hypernetworks for Generalizable BRDF Estimation, Fazilet Gokbudak et al., 2023.*

Contributions

Two new material datasets,

multi-modal synthesis of data-driven material appearance models,

a novel statistics-based constrained synthesis,

two BRDF distributional evaluation metrics,

further ablation studies are conducted.

Future work

developing physically accurate neural representations of BRDFs,

extending the approach to support more complex materials,

detailed statistical evaluation of multi-modality synthesis,

...

M3ashy: Multi-Modal Material Synthesis via Hyperdiffusion

https://peterhuistyping.github.io/M3ashy/

Thank you for listening ~

Welcome for any questions or comments!