UMat: Uncertainty-Aware Single Image High Resolution Material Capture

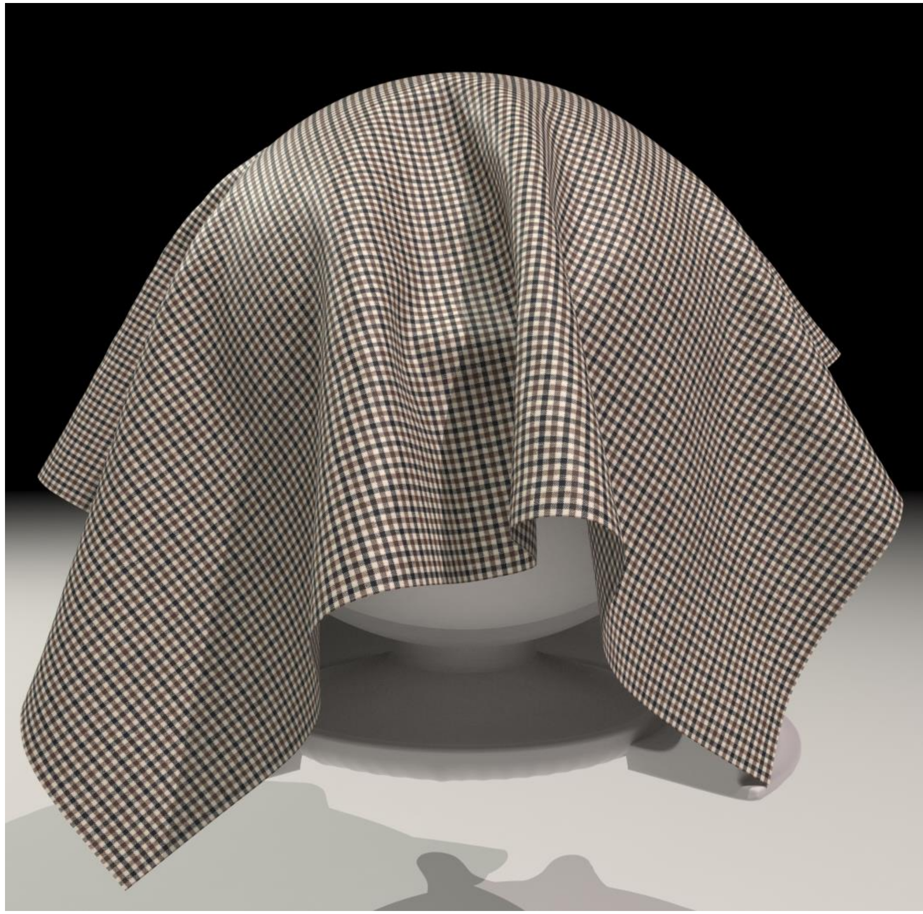
CVPR 2023
Paper ID 7763
TUE-PM-156

SEDDI

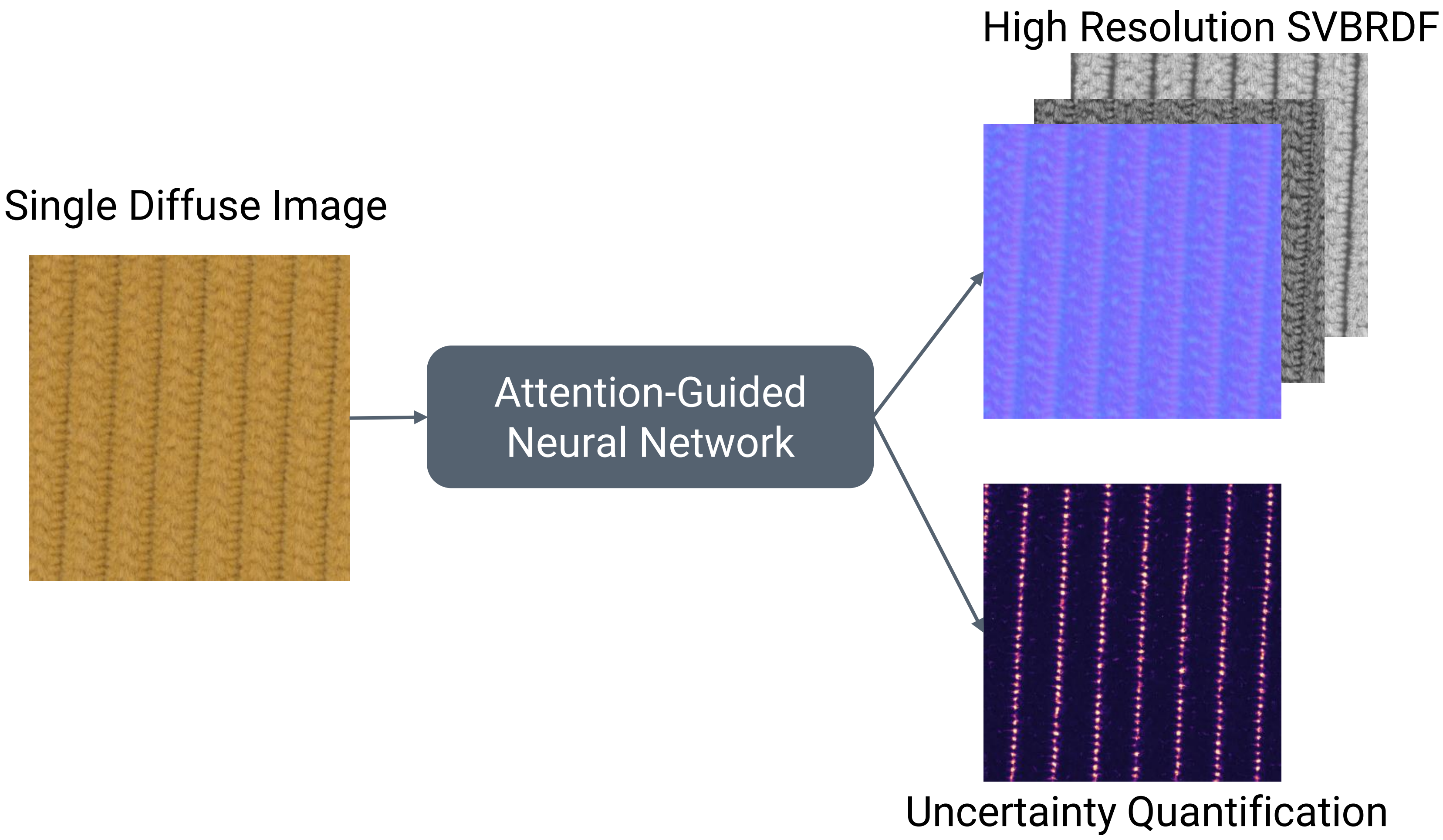
Universidad Rey Juan Carlos

1-Minute Summary

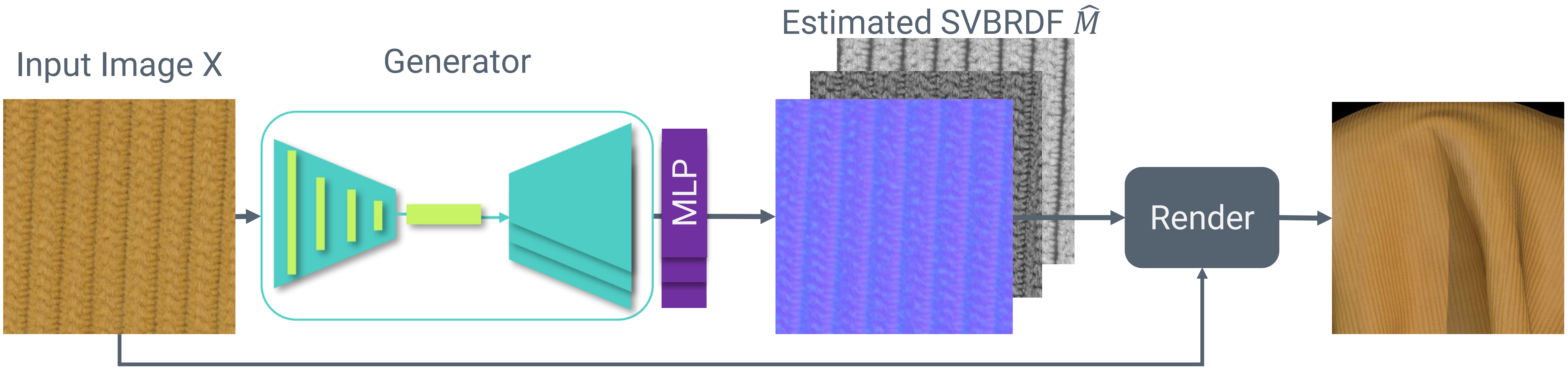
Introduction



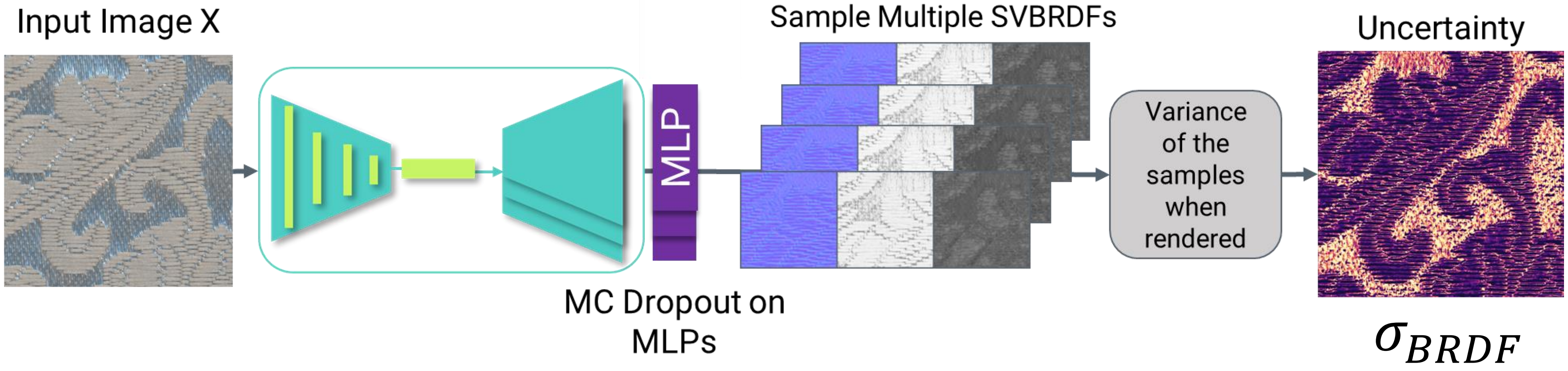
UMat

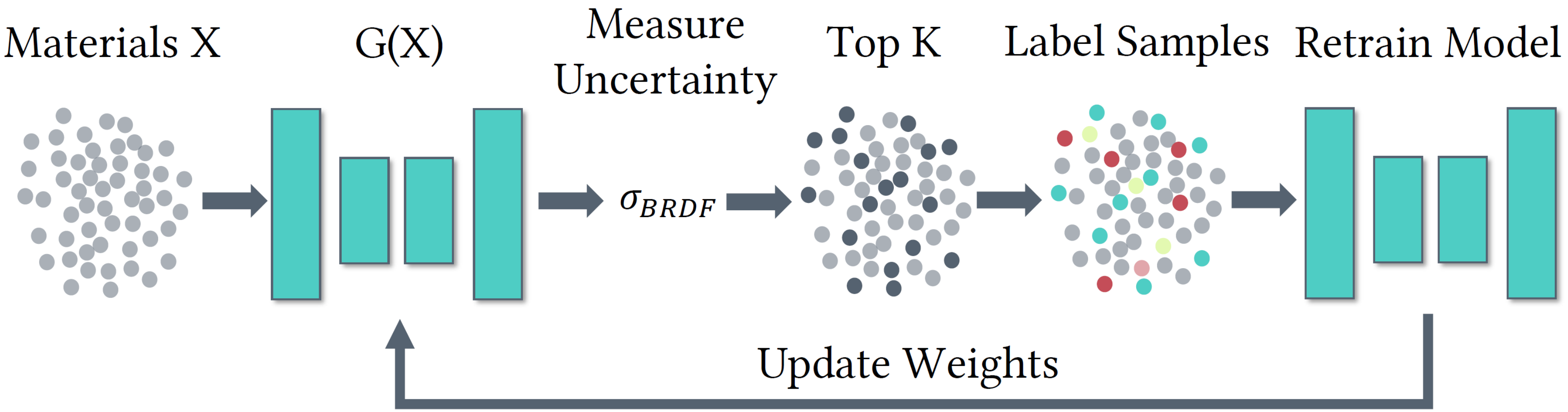

High Resolution SVBRDF
Single Diffuse Image
Attention-Guided
Neural Network
Uncertainty Quantification

Inference


Input Image X
Generator
MLP
Estimated SVBRDF $\widehat{M}$
Render

Uncertainty Quantification and Active Learning





UMat: Uncertainty-Aware Single Image High Resolution Material Capture

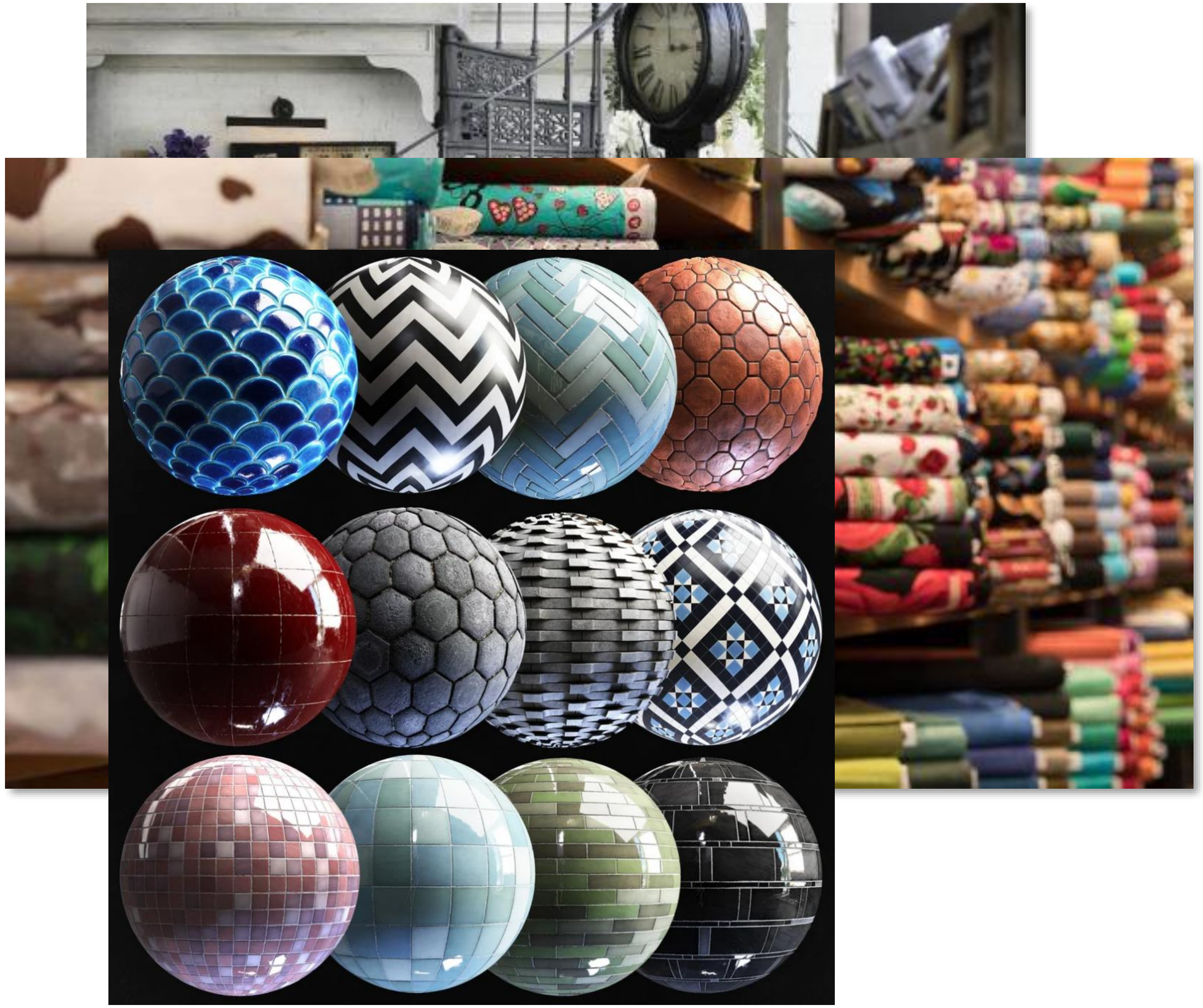
CVPR 2023
Paper ID 7763
TUE-PM-156

SEDDI

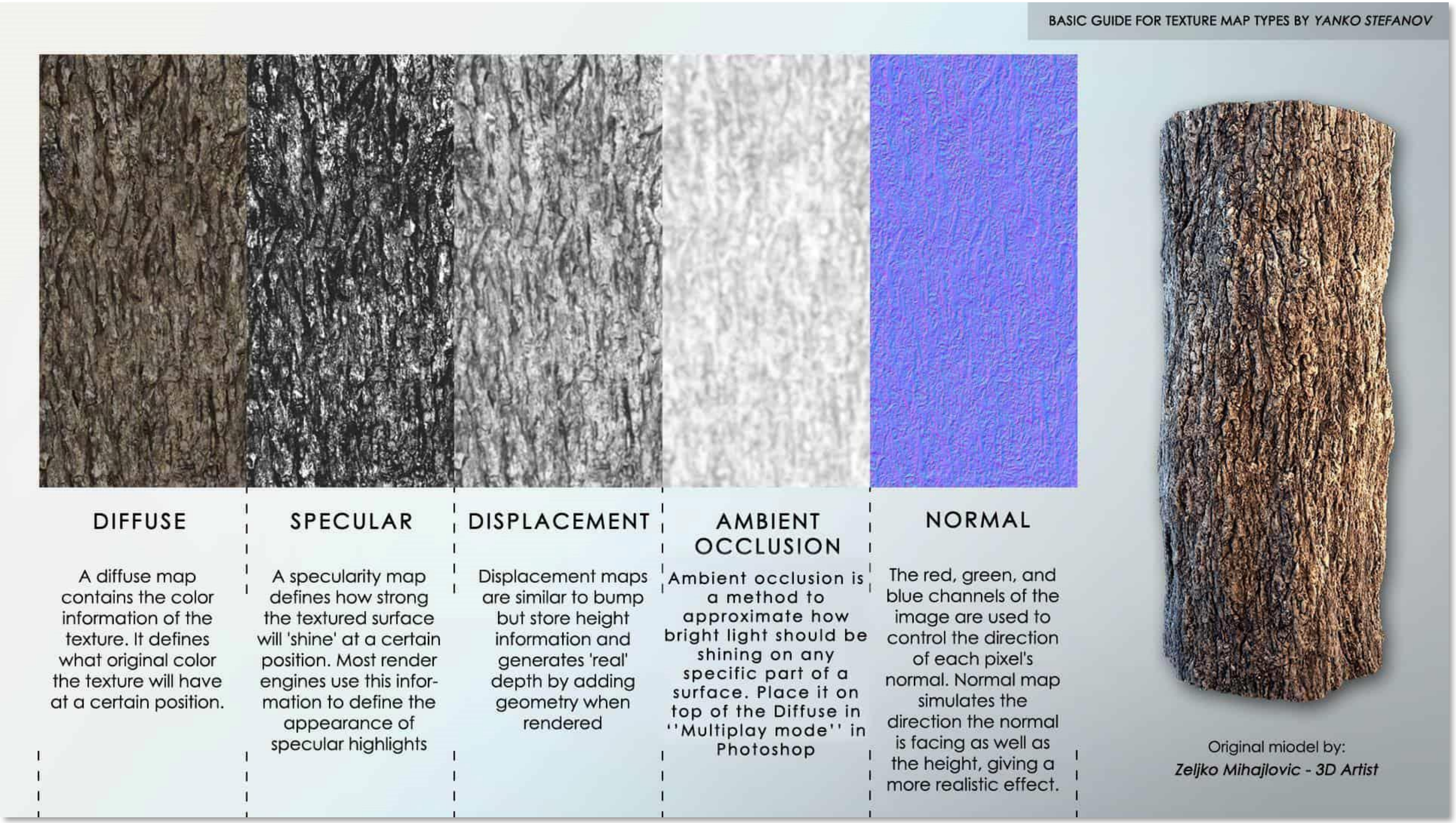
Universidad Rey Juan Carlos

Introduction

Real World Materials

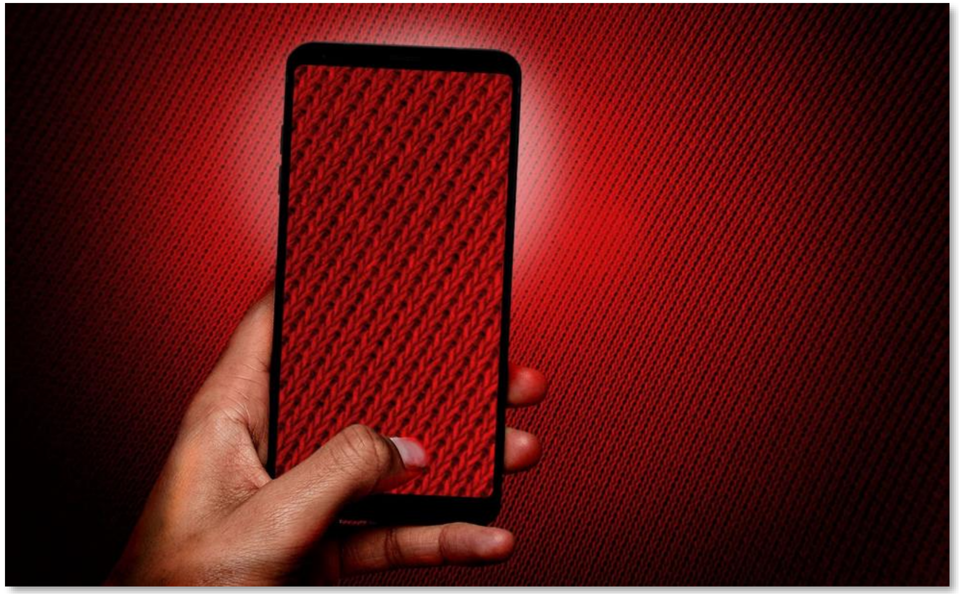


SVBRDFs


BASIC GUIDE FOR TEXTURE MAP TYPES BY YANKO STEFANOV
DIFFUSE
A diffuse map contains the color information of the texture. It defines what original color the texture will have at a certain position.
SPECULAR
A specularity map defines how strong the textured surface will 'shine' at a certain position. Most render engines use this information to define the appearance of specular highlights
DISPLACEMENT
Displacement maps are similar to bump but store height information and generates 'real' depth by adding geometry when rendered
AMBIENT OCCLUSION
Ambient occlusion is a method to approximate how bright light should be shining on any specific part of a surface. Place it on top of the Diffuse in ''Multiplay mode'' in Photoshop
NORMAL
The red, green, and blue channels of the image are used to control the direction of each pixel's normal. Normal map simulates the direction the normal is facing as well as the height, giving a more realistic effect.
Original miodel by:
Zeljko Mihajlovic - 3D Artist

Material Capture

Commodity Hardware



Custom Devices (Gonio Reflectometers, TAC7)

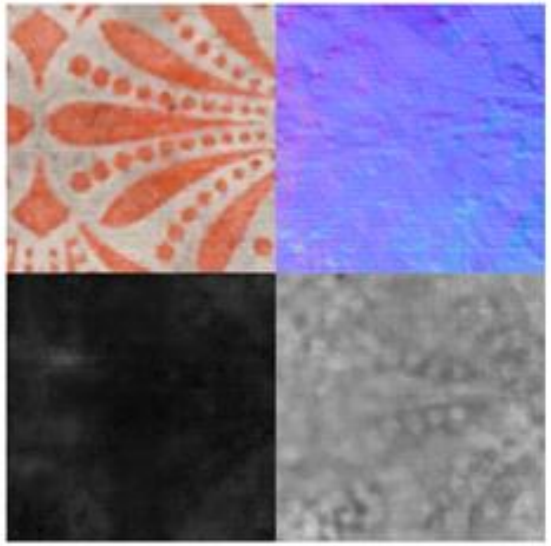
← Scalability, Speed

Cost, Quality →

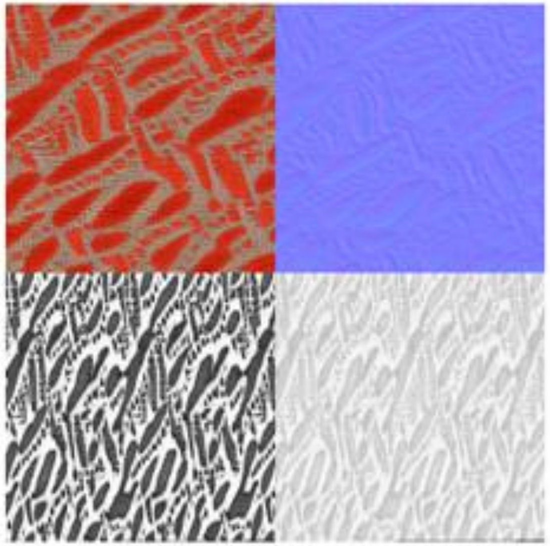
Material Capture

Smartphone Image

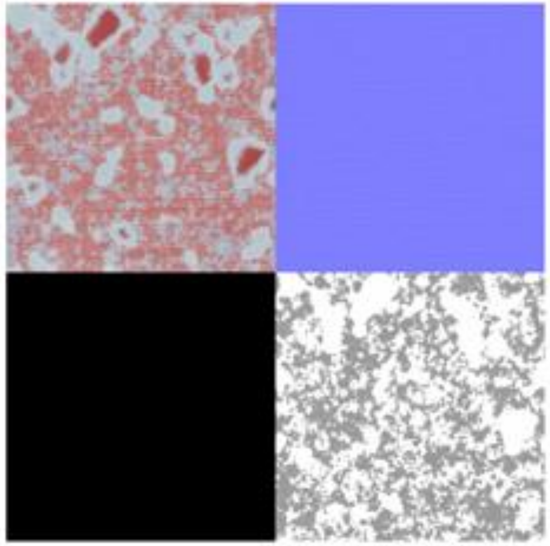
Gao et al.
2019



Henzler et al.
2021



Shi et al.
2020



UMat

GAN tailored for material digitization

High Resolution SVBRDF (up to 1000 ppi)

Flatbed Scanners as Capture Device

Accurate, artifact-free, sharp maps

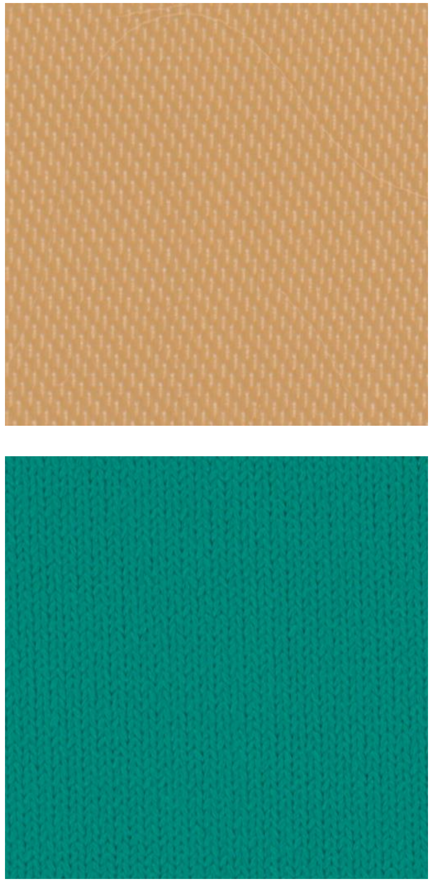
Uncertainty Quantification

Uncertainty in Single Image Material Estimation

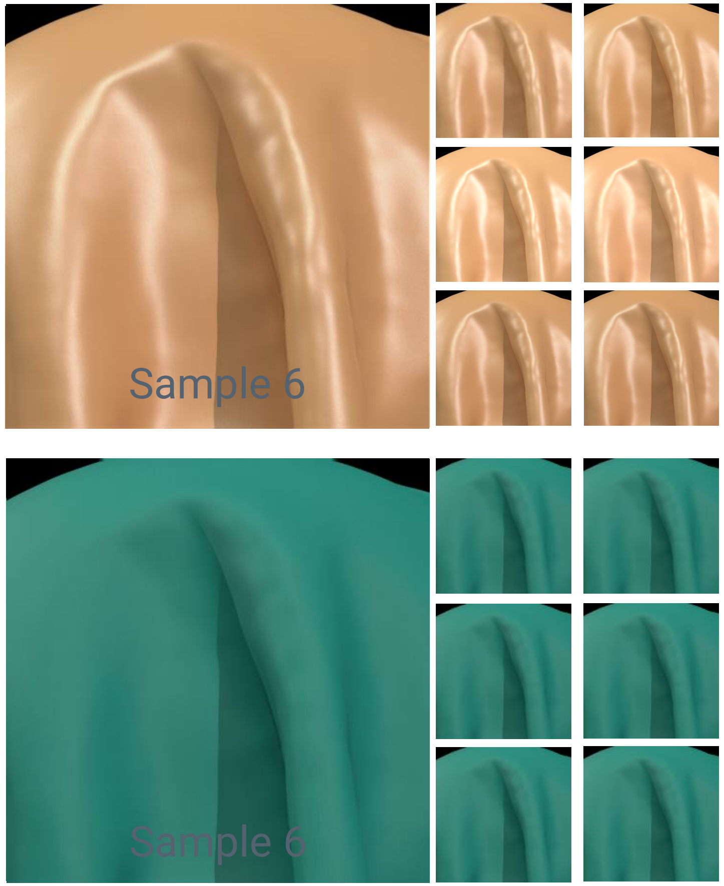
How will this look like?

Single Image Input

Single Image Input

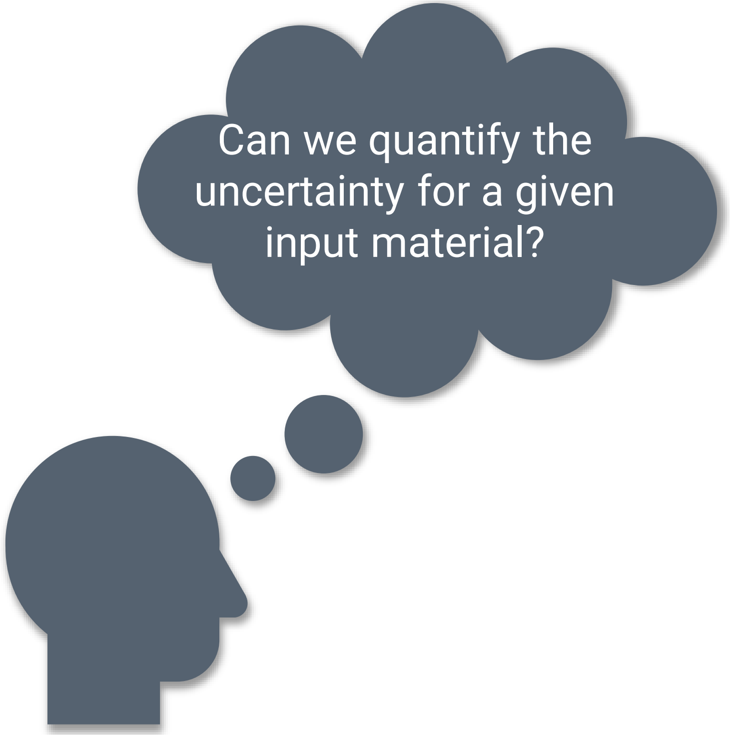


Plausible Model Estimations

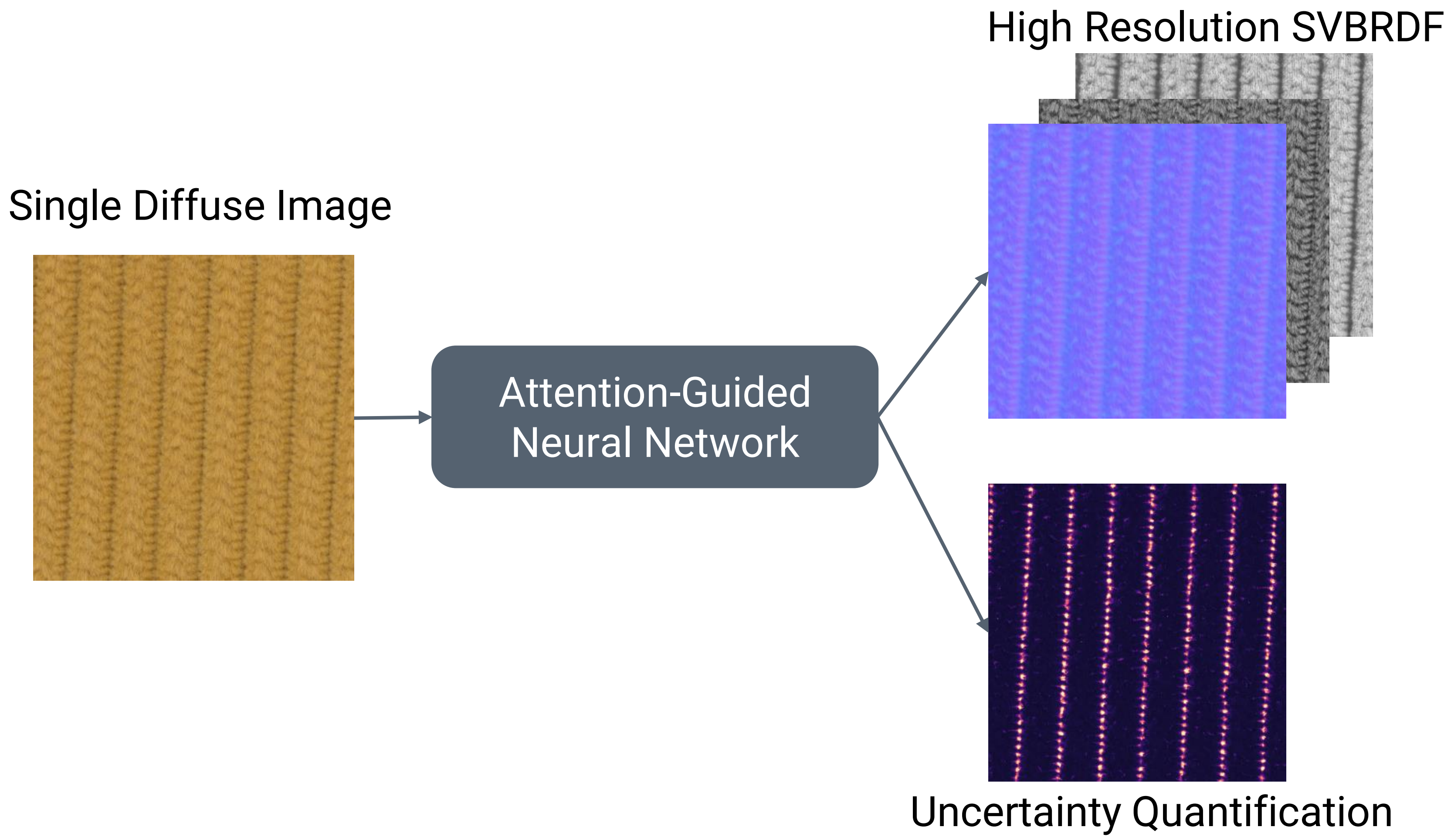


High Uncertainty

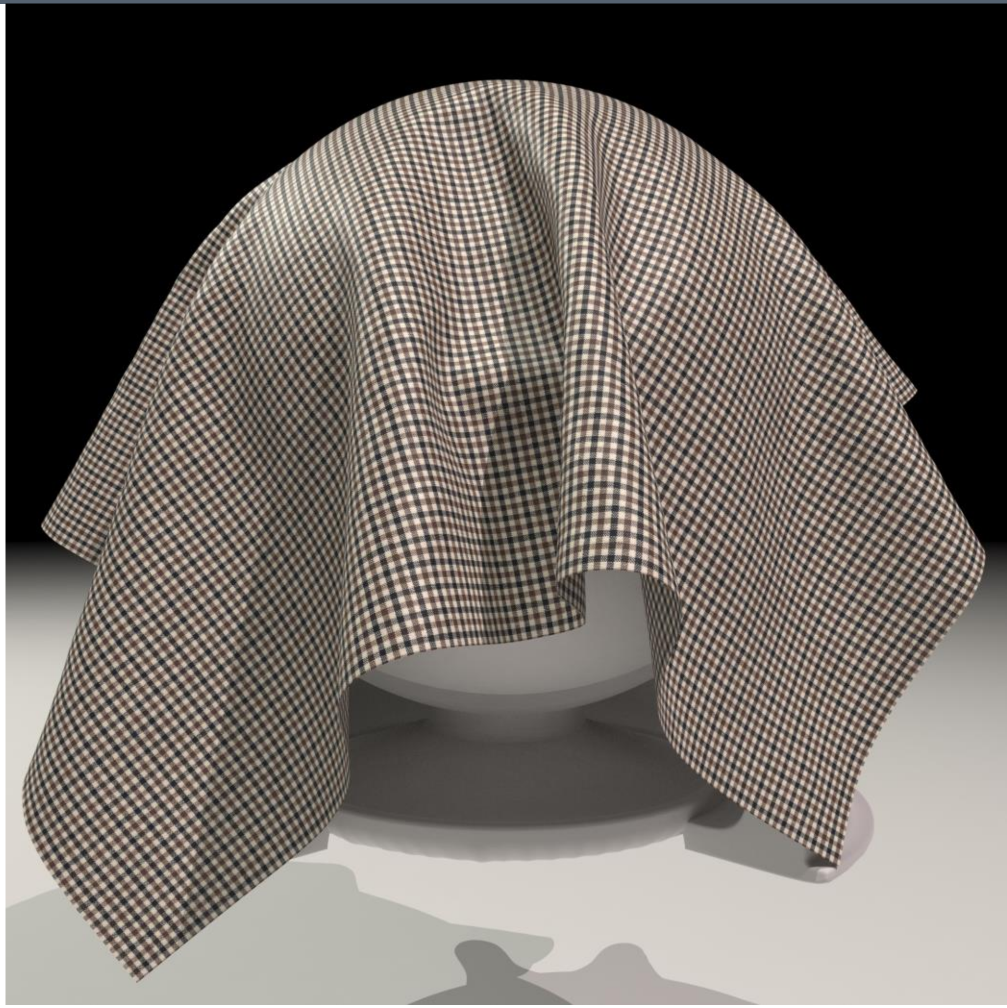
Low Uncertainty



UMat


Single Diffuse Image
Attention-Guided
Neural Network
High Resolution SVBRDF
Uncertainty Quantification

Flatbed Scanners



Flatbed Scanners

Scan

Albedo

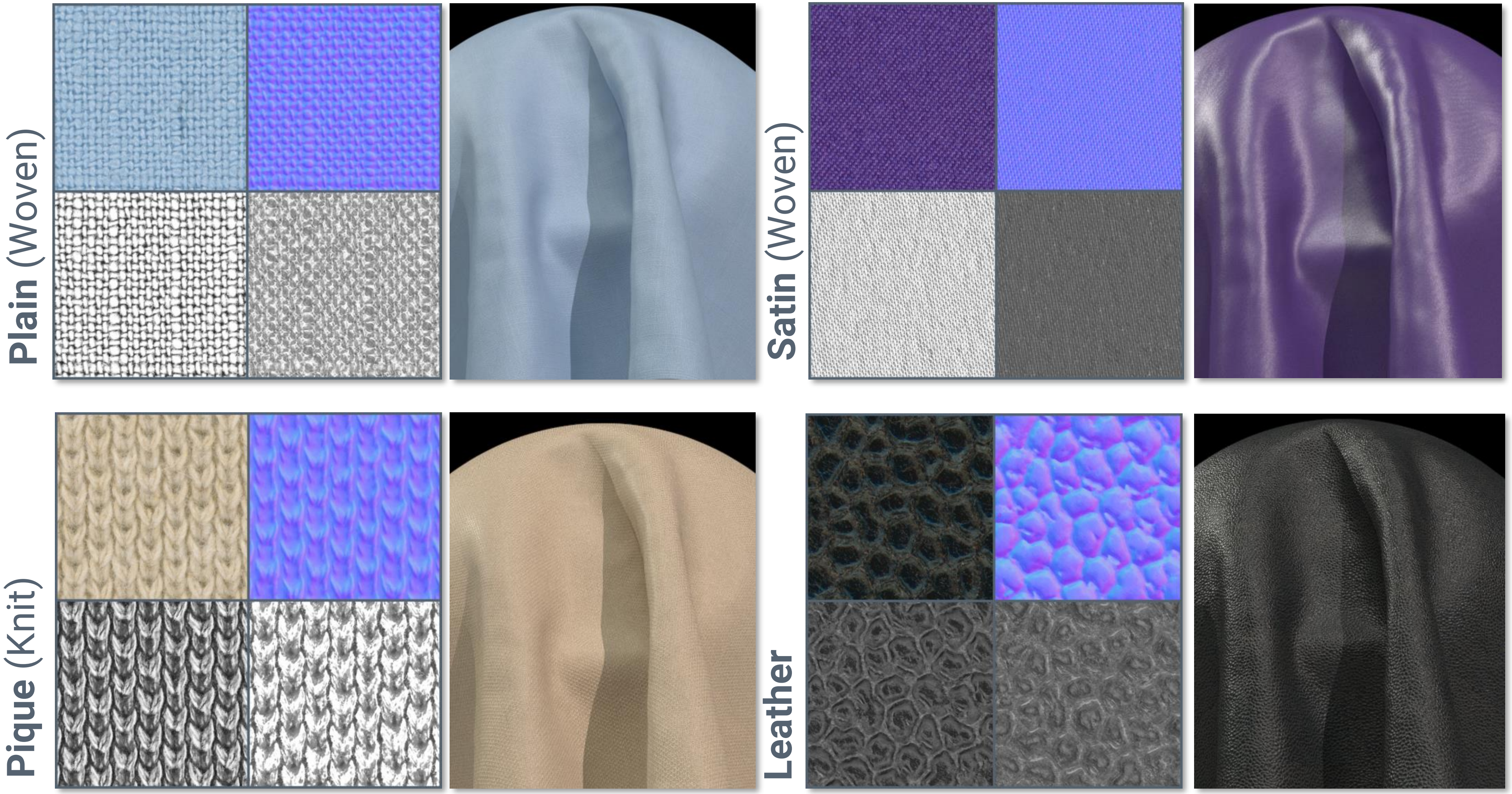
Scan

Albedo

Scan

Albedo

Estimating Reflectance from Microgeometry



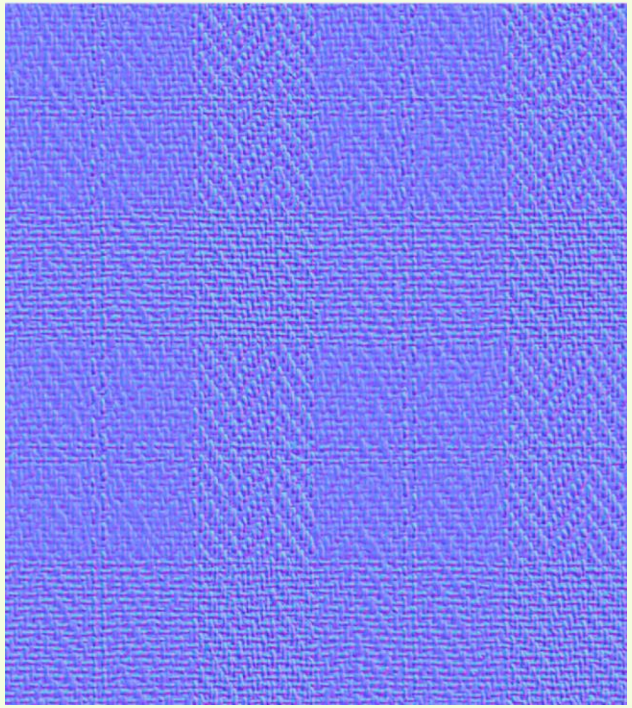
SVBRDF

SVBRDF

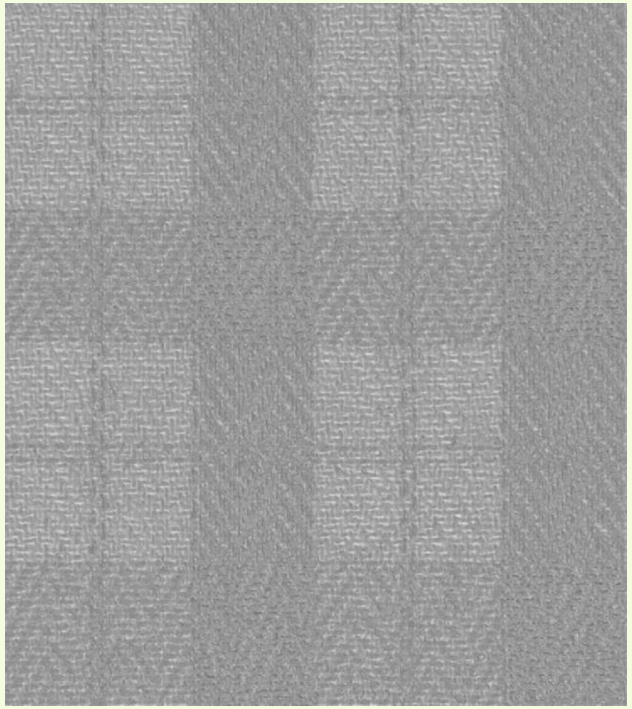
Input

Estimated

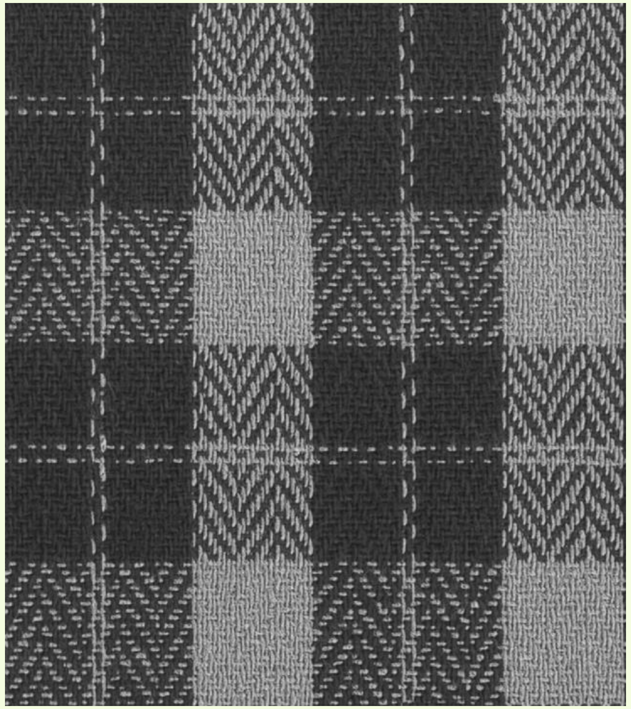
Albedo



Normals



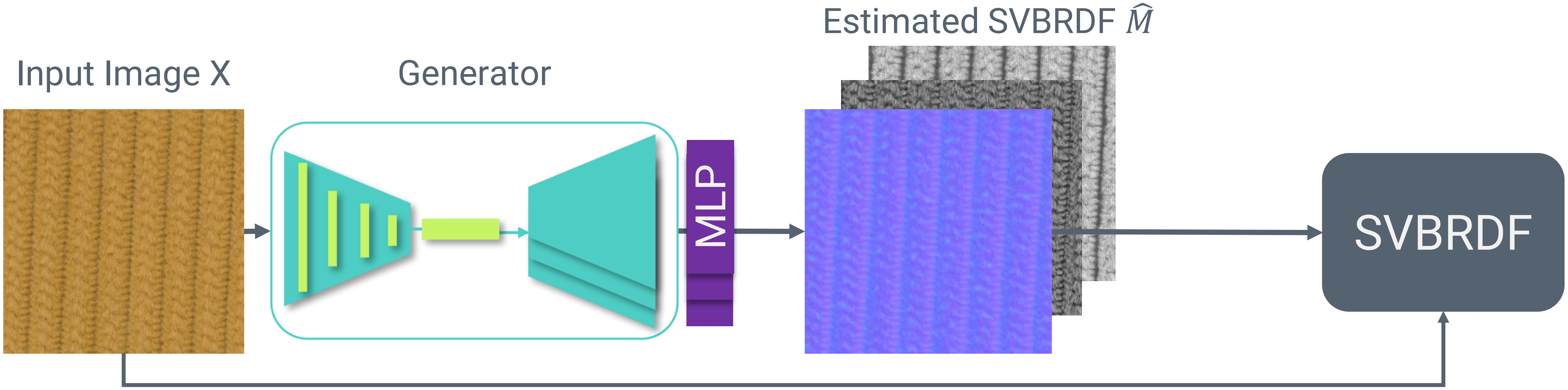
Roughness



Specular

SVBRDF

$$f_{l,v}(\mathbf{M}, \mathbf{X}) = \frac{\mathrm{X}}{\pi} + s_{l,v}(\mathbf{M}) \in \mathbb{R}^{x \times y}$$



X: Albedo (Input)

M (G(X)): Normals, Roughness, Specular

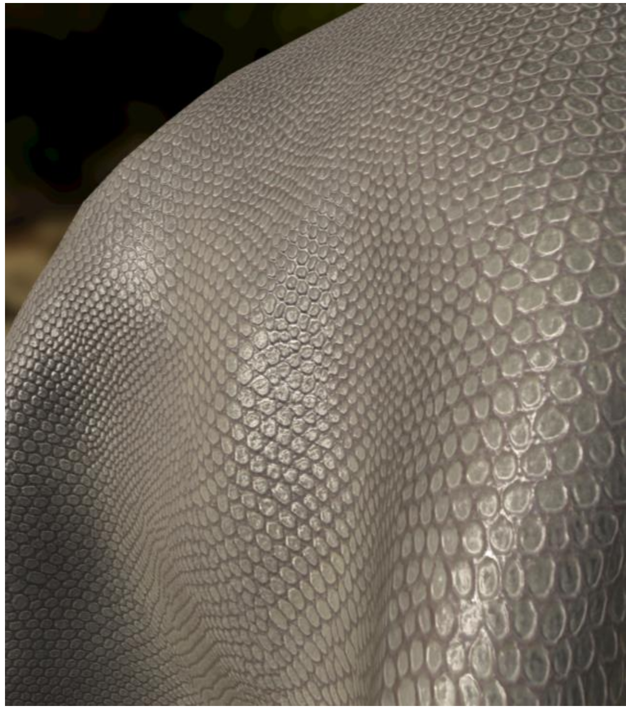
SVBRDF

Only Albedo

+Normals

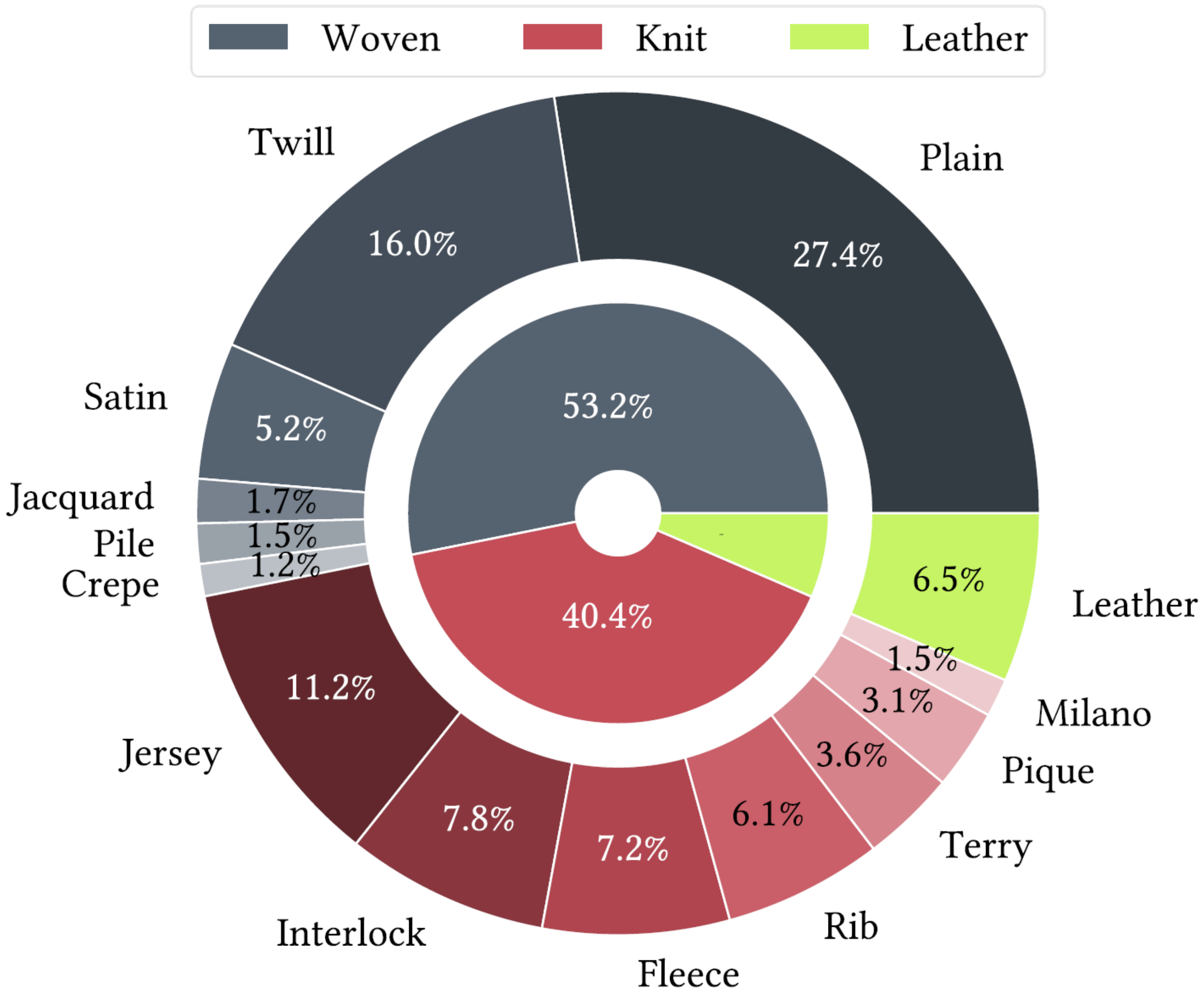
+Roughness

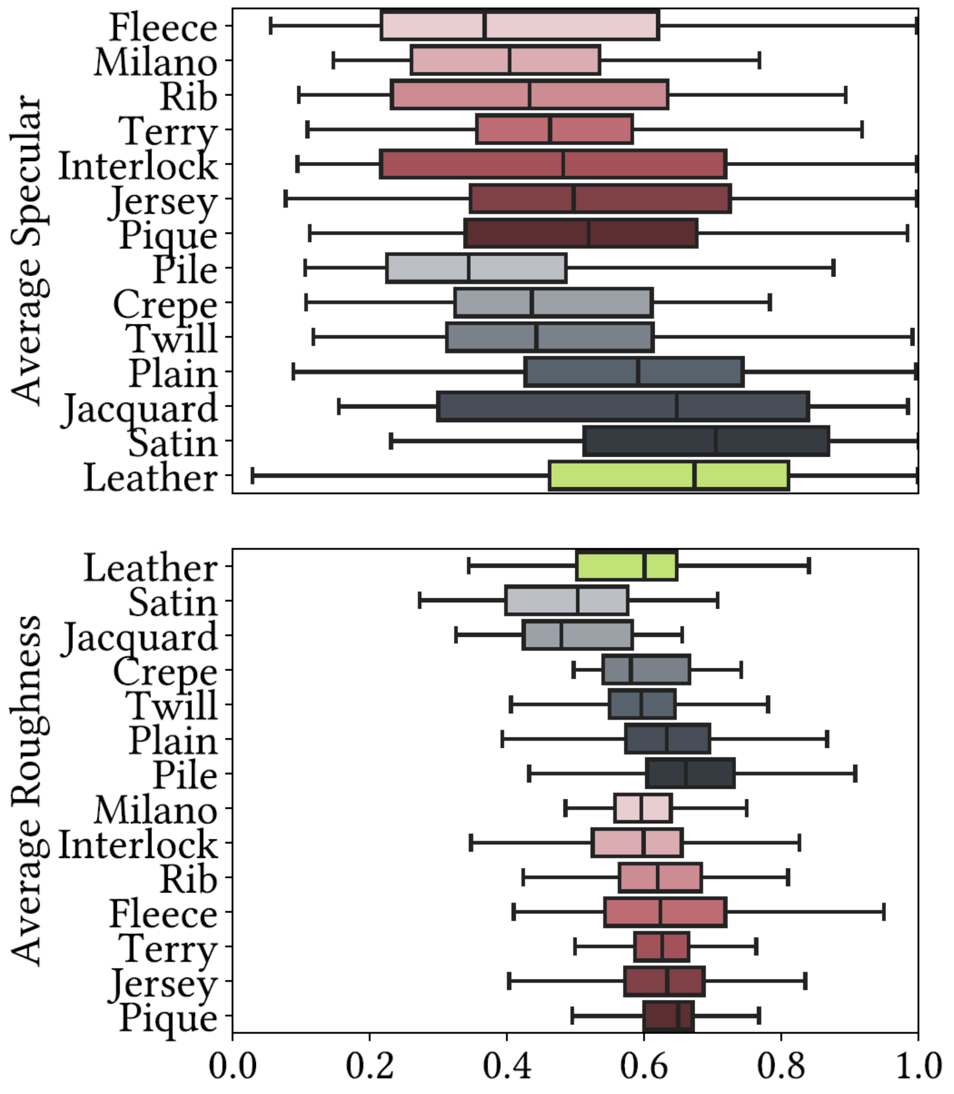
+Specular



Dataset

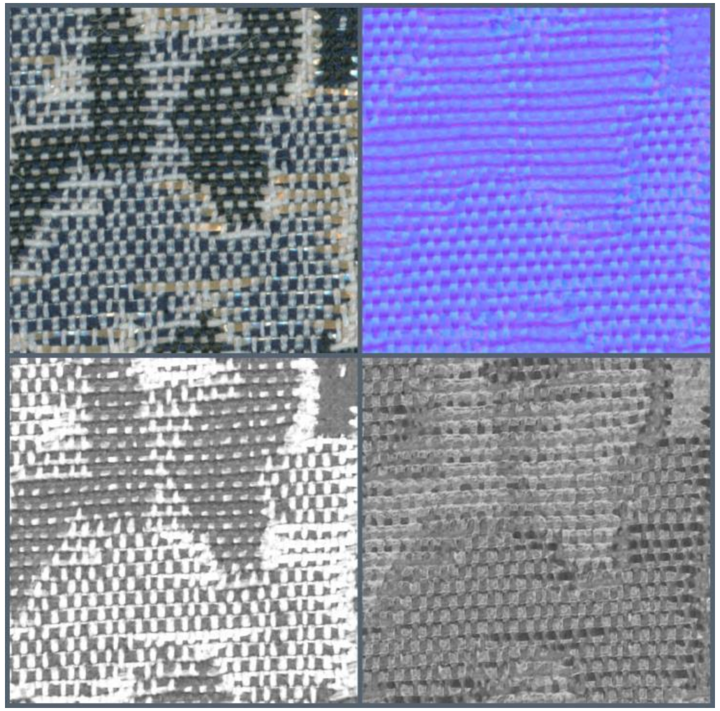
UMat: Dataset



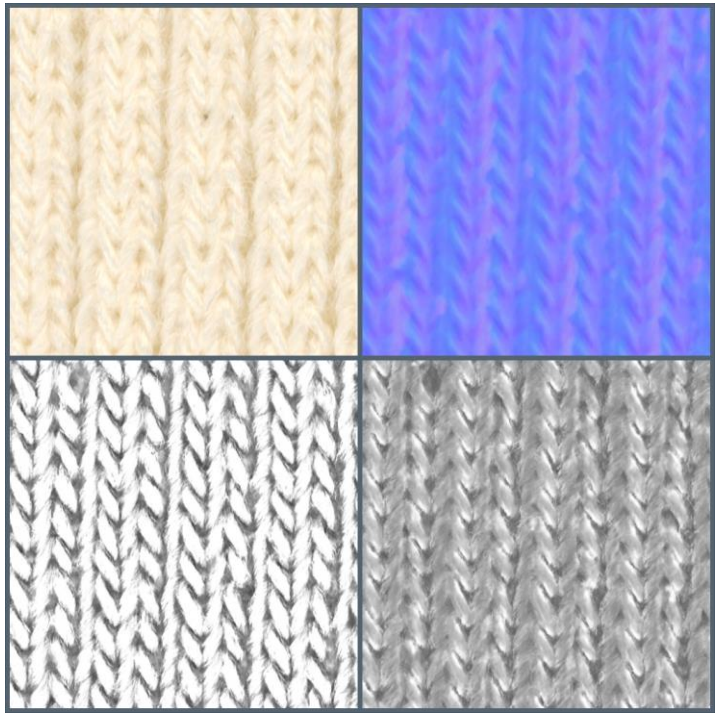


UMat: Dataset

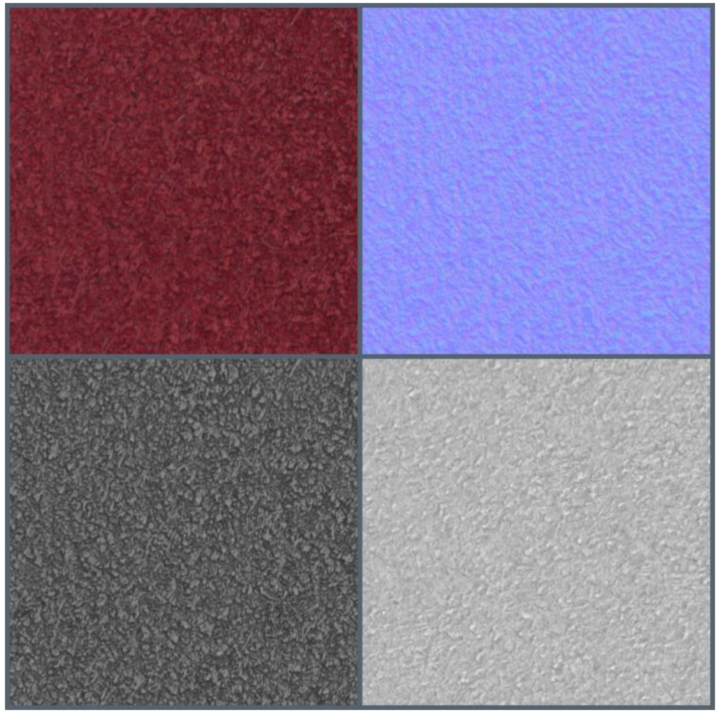
Jacquard (Woven)



Rib (Knit)



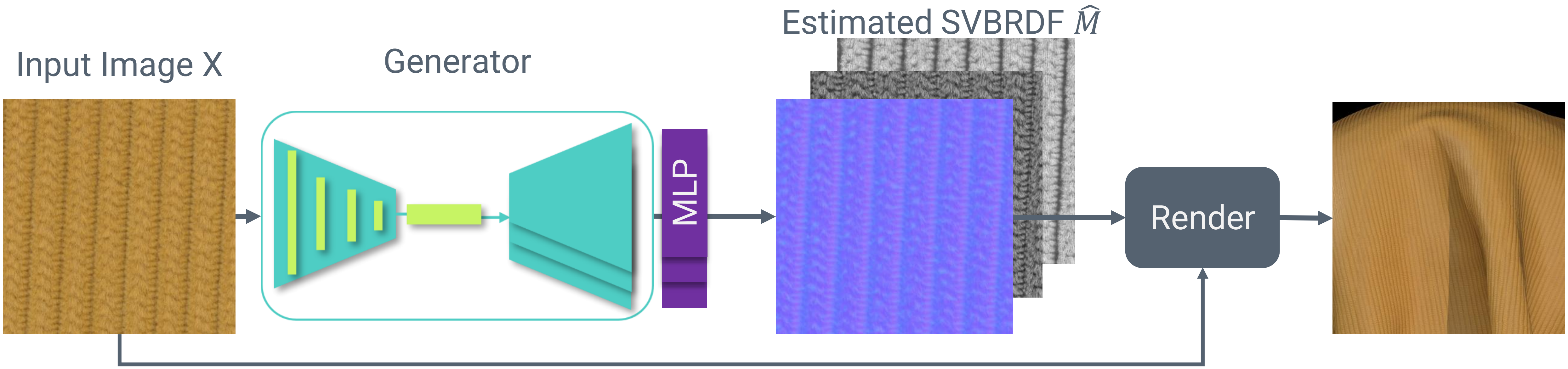
Suede (Leather)



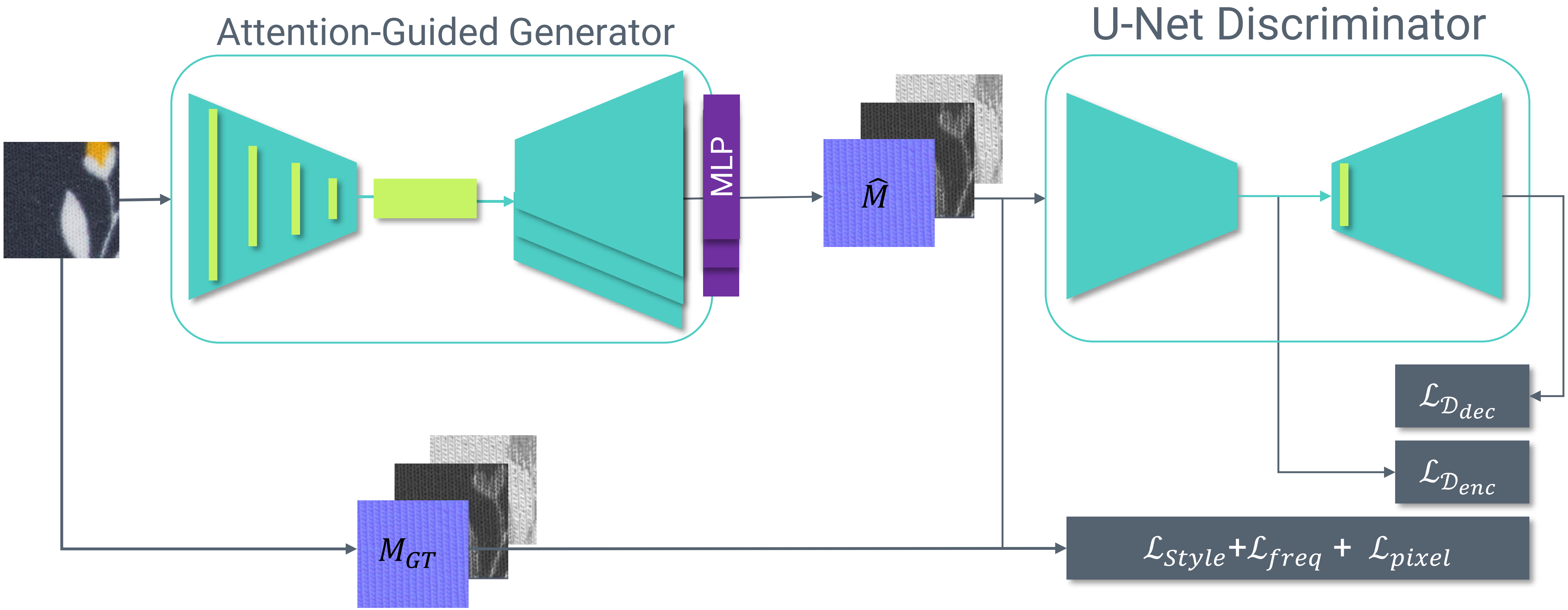
Towards Material Digitization with a Dual-scale Optical System.
Garces et al (TOG, Proc. SIGGRAPH 2023).

Model Design and Training

Inference

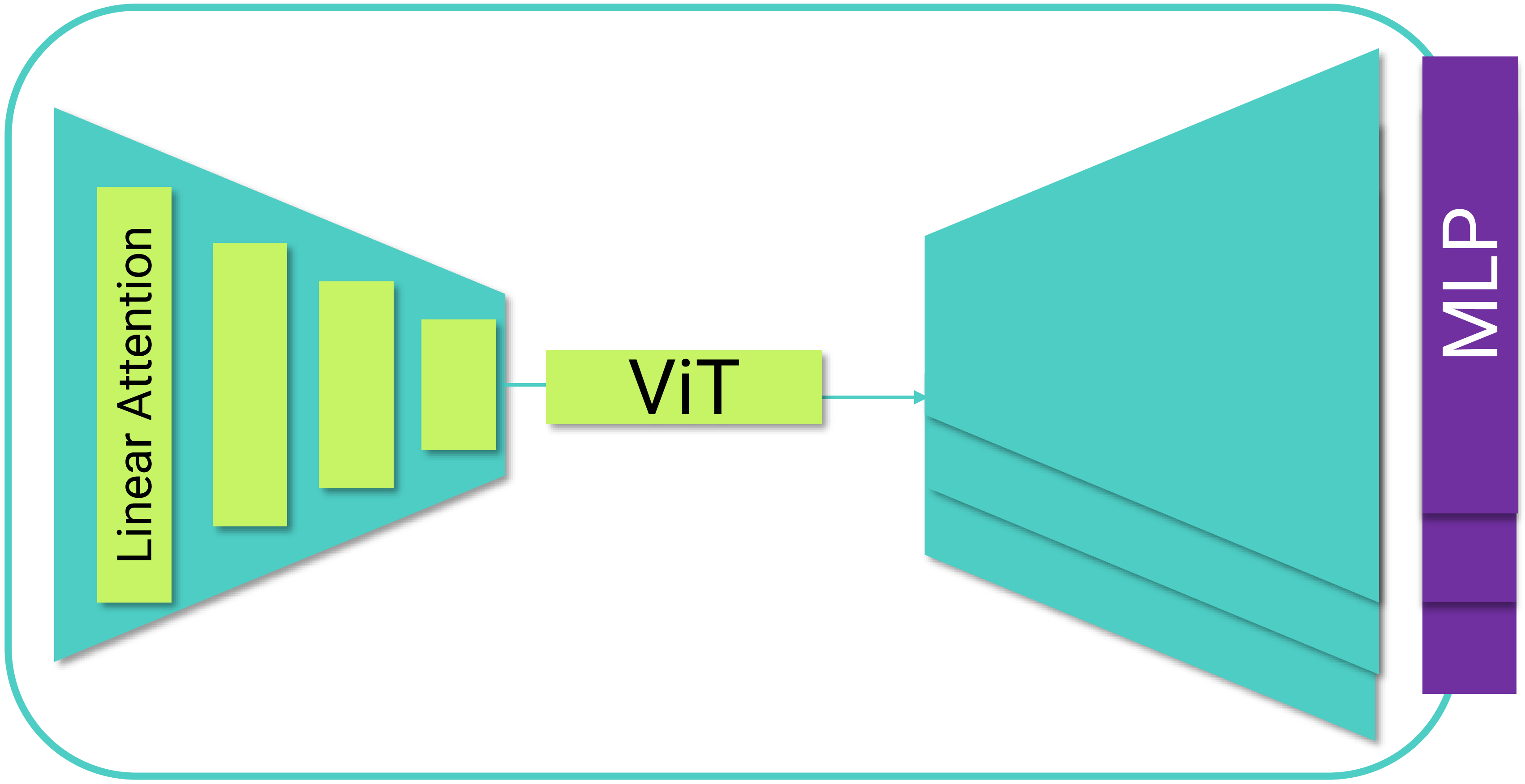

Input Image X
Generator
MLP
Estimated SVBRDF $\widehat{M}$
Render

Training



Attention-Guided Generator
MLP
$\widehat{M}$
U-Net Discriminator
$\mathcal{L}_{D_{dec}}$
$\mathcal{L}_{D_{enc}}$
M_{GT}
$\mathcal{L}_{Style}+\mathcal{L}_{freq} + \mathcal{L}_{pixel}$

Generator Architecture



Losses

$$\mathcal{L}_G = \boxed{\sum_i \lambda_i \mathcal{L}_{pixel_i}} + \boxed{\lambda_{adv}\mathcal{L}_{adv}} + \boxed{\lambda_{style}\mathcal{L}_{style}} + \boxed{\lambda_{freq}\mathcal{L}_{freq}}$$

$$\boxed{\begin{aligned} \mathcal{L}_{\mathcal{D}} &= \mathcal{L}_{\mathcal{D}_{enc}} + \mathcal{L}_{\mathcal{D}_{dec}} + \lambda_{cons}\mathcal{L}^{cons}_{\mathcal{D}_{dec}} \\ \mathcal{L}_{adv} &= \log(\mathcal{D}_{enc}(\mathrm{G}(\mathrm{X})) + \log(\mathcal{D}_{dec}(\mathrm{G}(\mathrm{X})) \end{aligned}}$$

Pixel-wise norm for accurate maps

Style loss increases sharpness and perceptual quality

Pixel-wise and global adversarial losses for local and global quality

Frequency loss allows for better learning high-frequency patterns

Evaluation

BRDF Evaluation Error

$$\mathcal{L}_{\text{BRDF}} = \frac{1}{|xy|} \sum_{xy} \sqrt{\frac{1}{|S|} \sum_{(l,v) \in S} \boxed{\sqrt[3]{\cos^2(\theta_l)}} \left(f_{l,v}(\mathbf{M}_{GT}, K) - f_{l,v}(\hat{\mathbf{M}}, K) \right)^2}$$

Average render distance between M_{GT} and estimated $\widehat{M}$

We render $f_{l,v}$ the SVBRDFs at a set of optimized light and view positions

We use a grayscale albedo K to isolate the impact of the other parameters

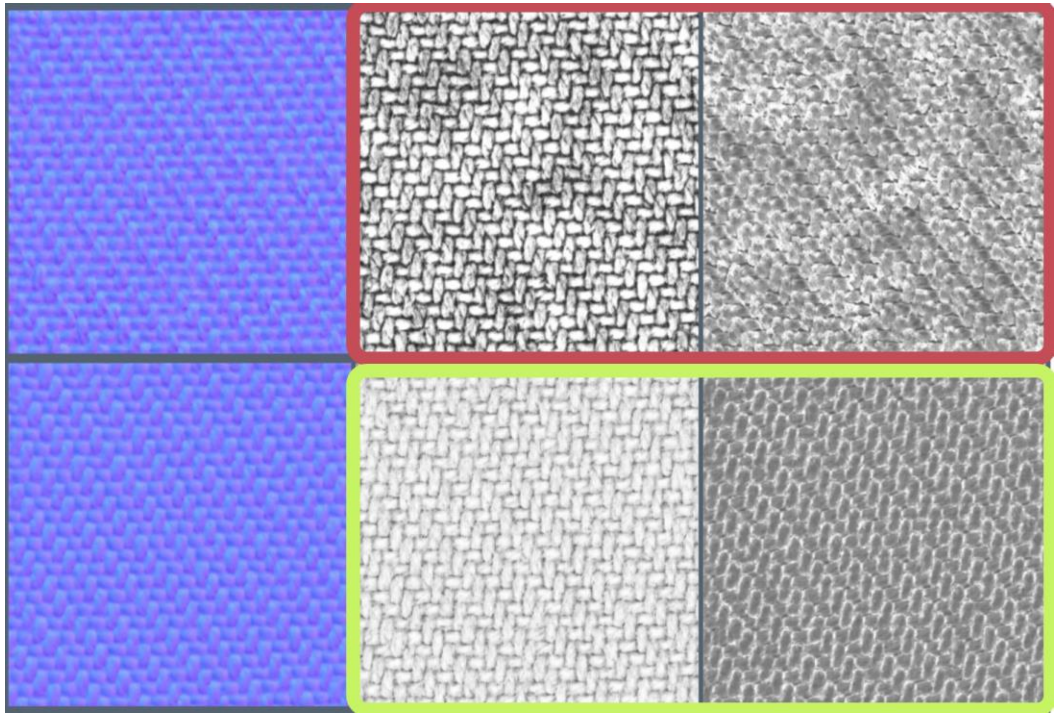
Perceptually Motivated: Specular Peak attenuation, Cosine Weighting

Artifact Detection

Input Image

Estimation with artifacts

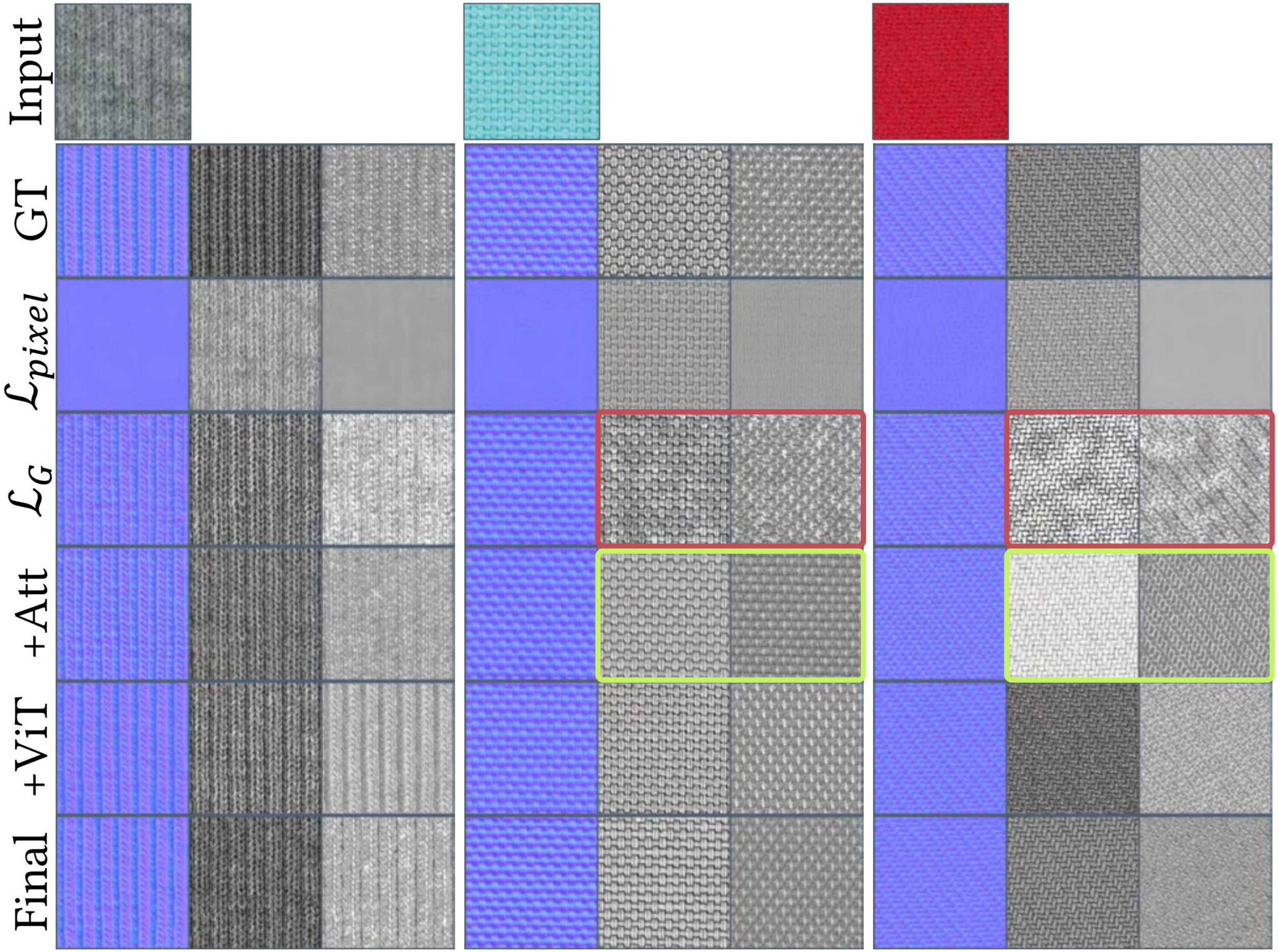
Artifact-free estimation



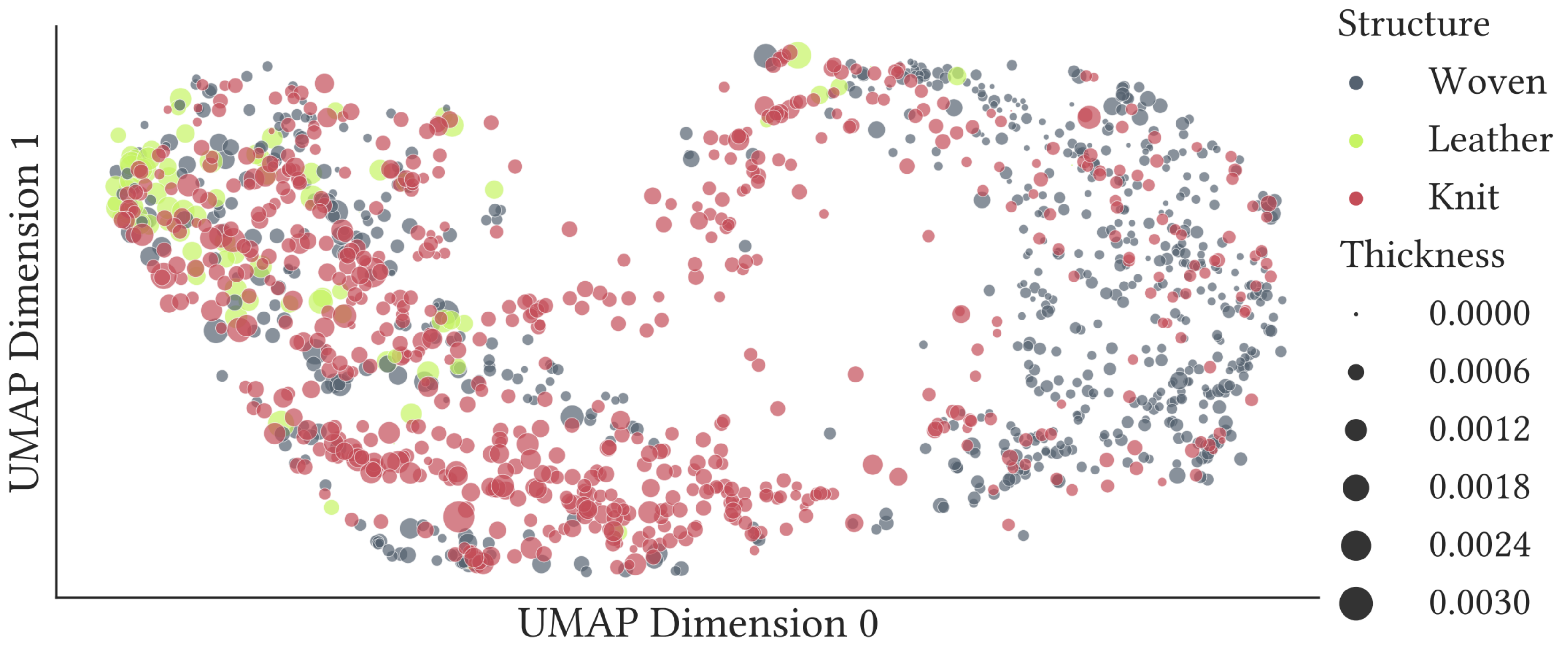
Mutual information for automatic artifact detection

$$\mathcal{H}(\mathbf{I}) = \frac{1}{|xy|} \sum_{xy} \frac{1}{|d|} \sum_{d=\{\uparrow,\downarrow,\leftarrow,\rightarrow\}} \| F_{\mathrm{Box}}(\mathbf{I}) - F_{\mathrm{Box}}(\mathbf{I}^{d}) \|_1$$

UMat: Qualitative Ablation

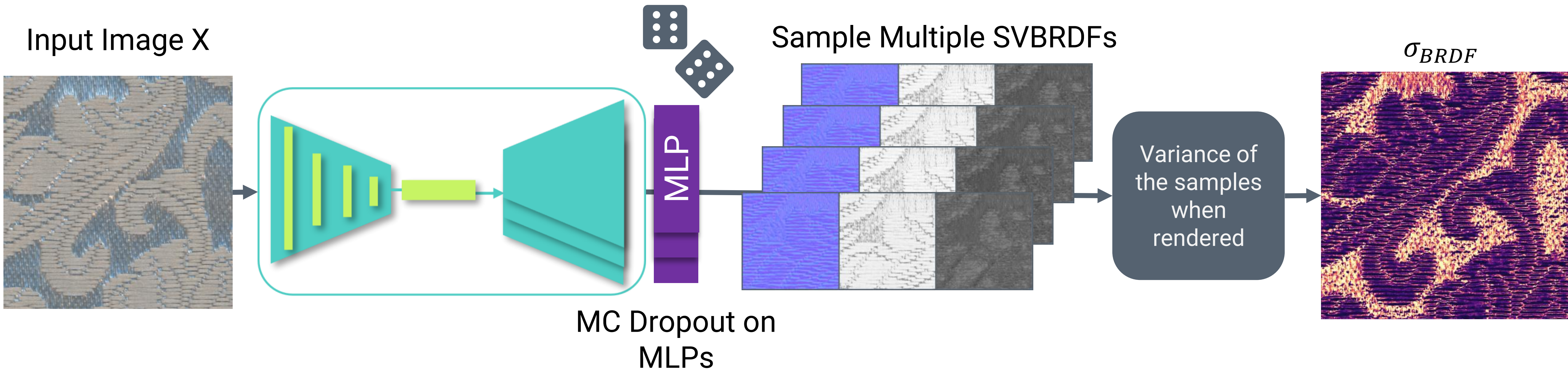


Latent Embeddings



Uncertainty Quantification

Uncertainty Estimation: Our Approach



Uncertainty Estimation: Our Approach

$$\sigma_{\text{BRDF}} = \frac{1}{|xy|}\sum_{xy}\log\left(\frac{1}{|S|}\sqrt{\sum_{(l,v)\in S}\boxed{\sqrt[3]{\sigma_{l,v}}}\left(\{f_{l,v}\boxed{(\mathrm{U}_j,}K)\boxed{\cos(\theta_l)\}_{j=1}^{N})}}\right)$$

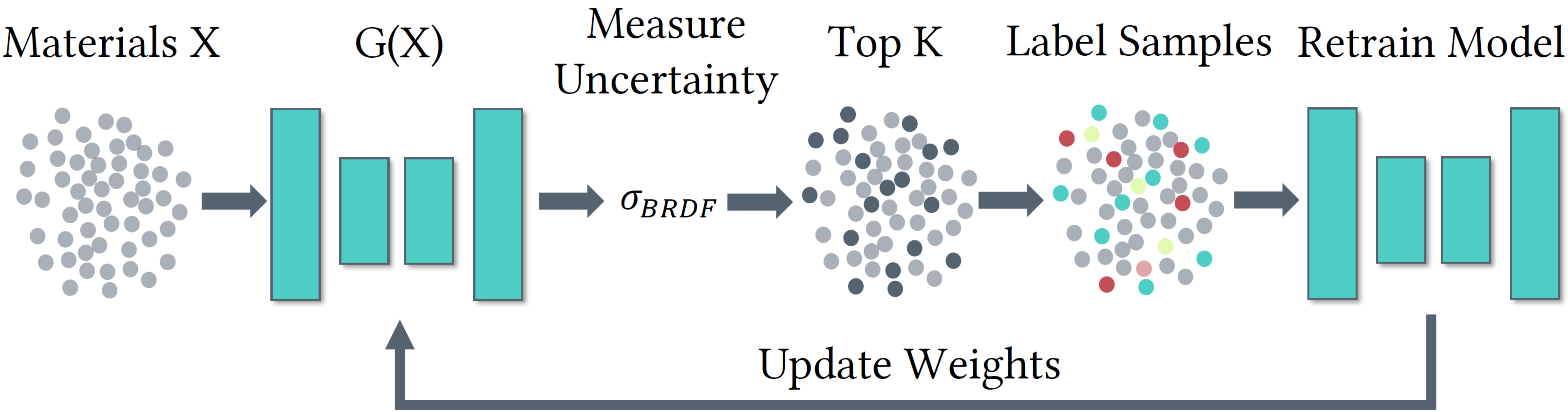
We sample a set (**U**) of estimations using **Monte Carlo Dropout** on the MLPs.

We render $f_{l,v}$ using a grayscale albedo K

Perceptually Motivated: Specular Peak attenuation, Cosine Weighting

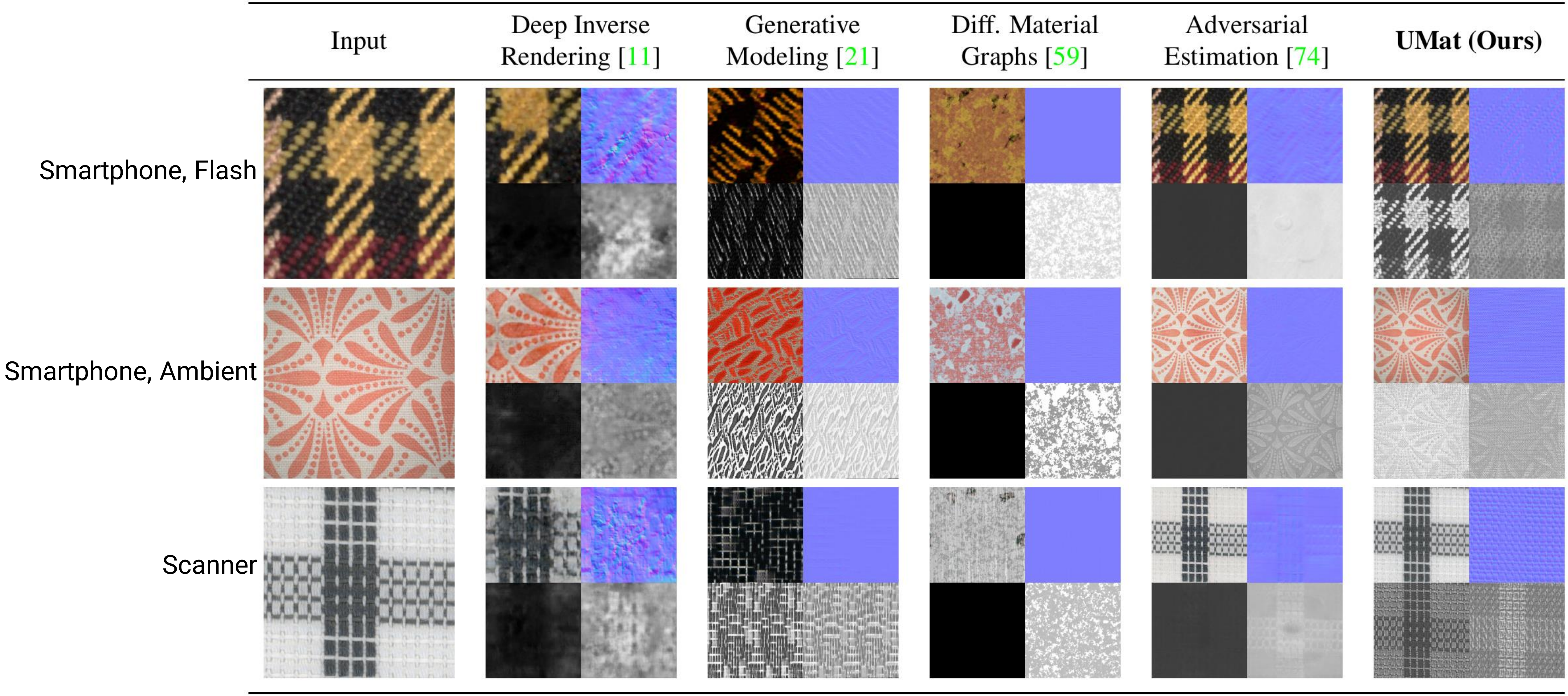
Variance across renders generated at a set of light (**l**) and cameras (**v**) positions

Application: Active Learning

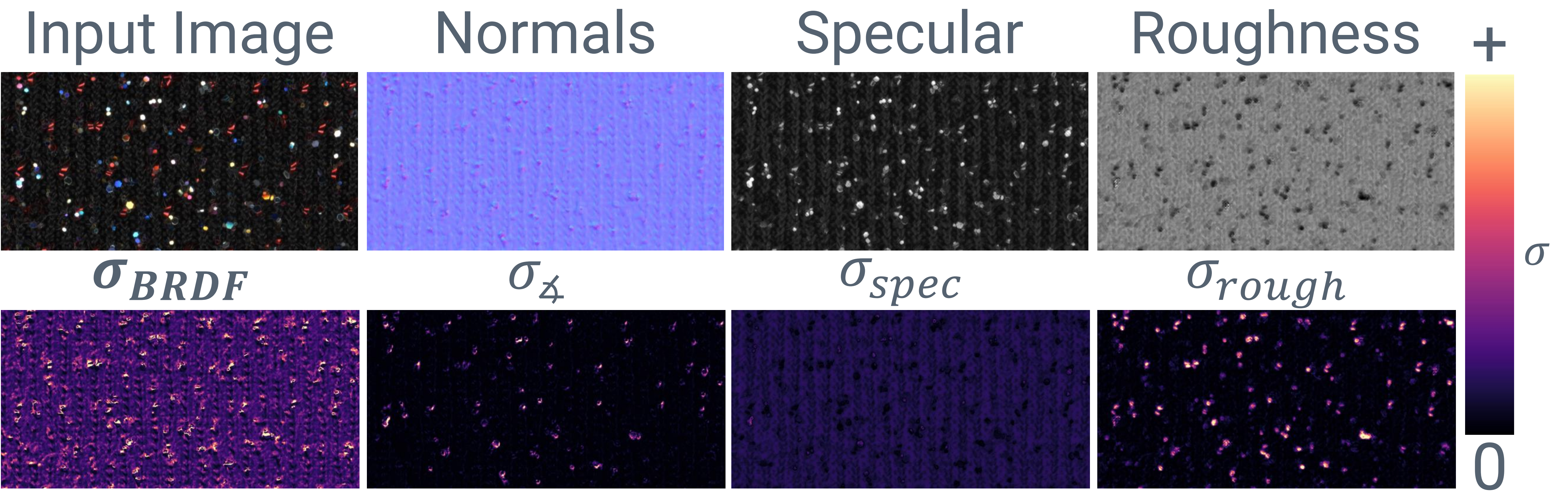


Results

UMat: Comparisons with Previous Work

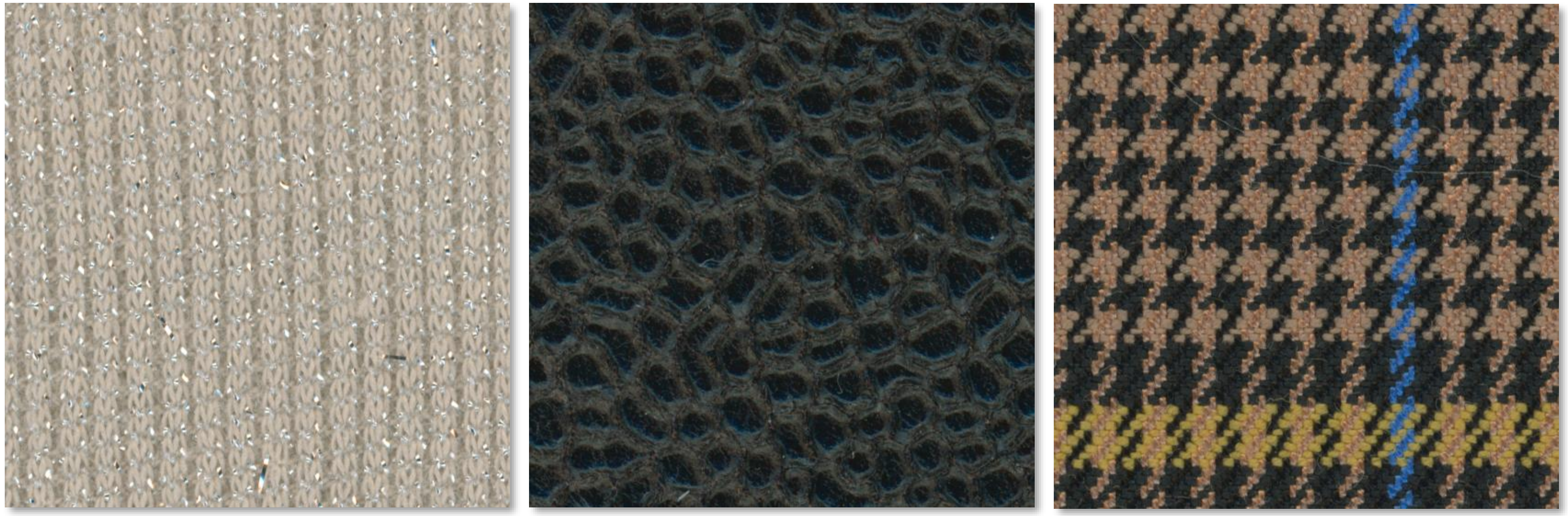


Uncertainty Quantification: Results

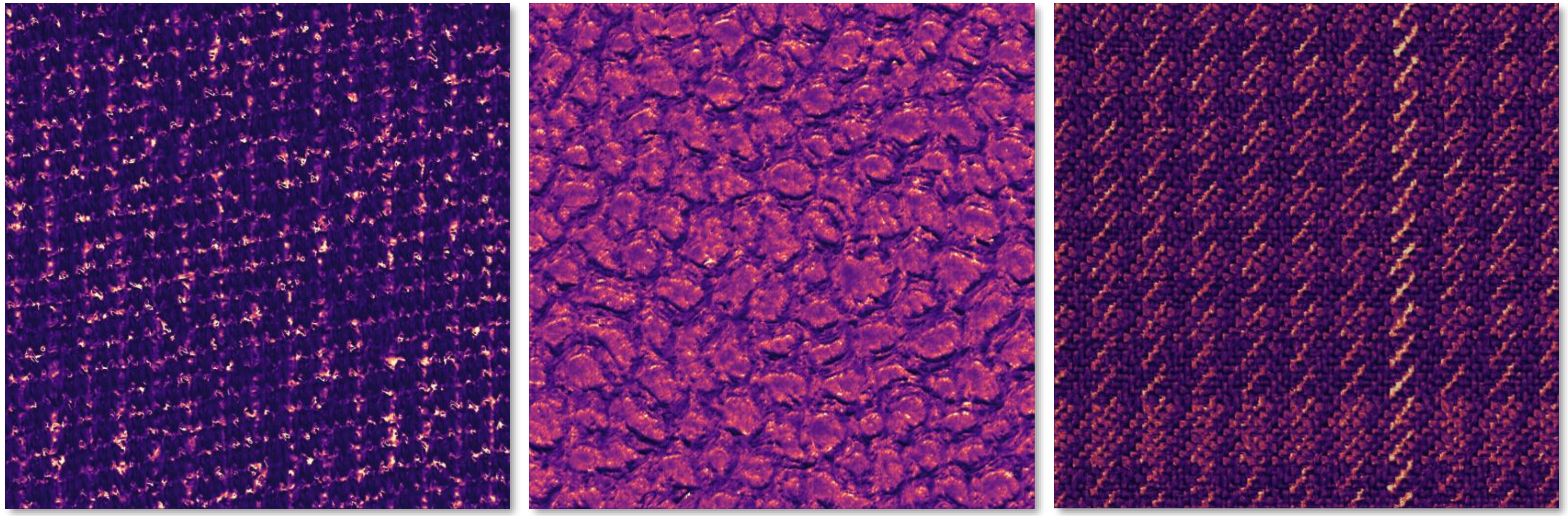


Uncertainty Quantification: Results

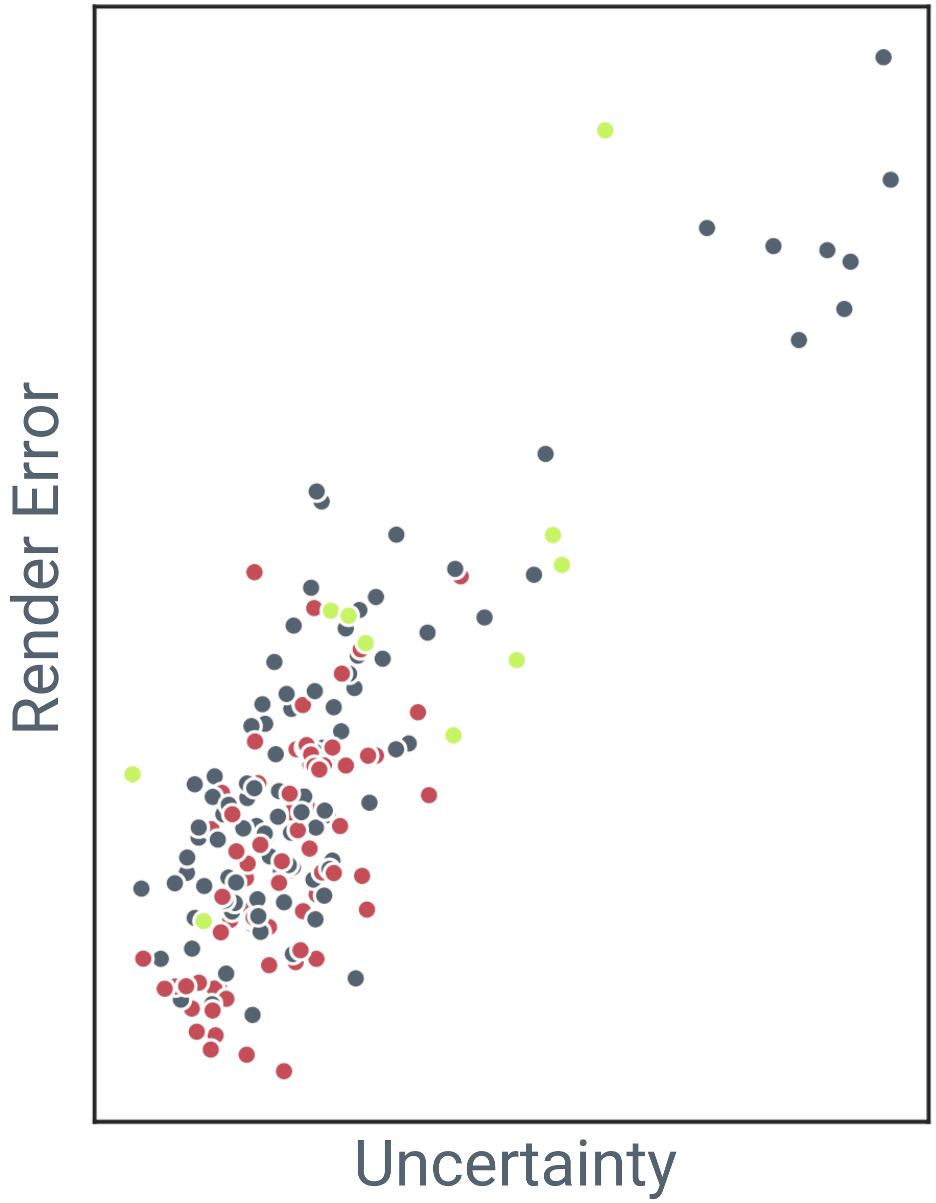
Metallic yarns | Leather | Tartan



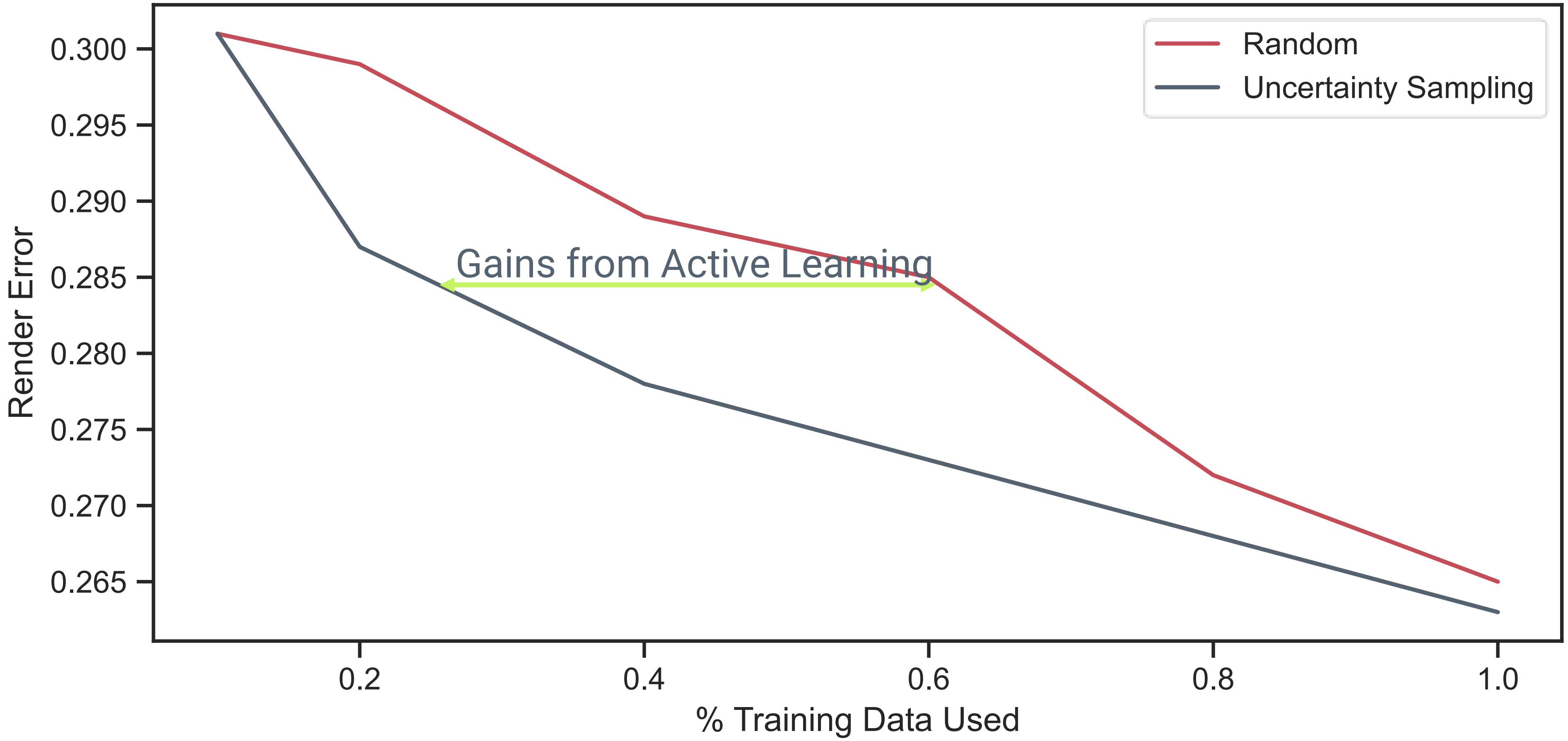
σ_{BRDF} | σ_{BRDF} | σ_{BRDF}



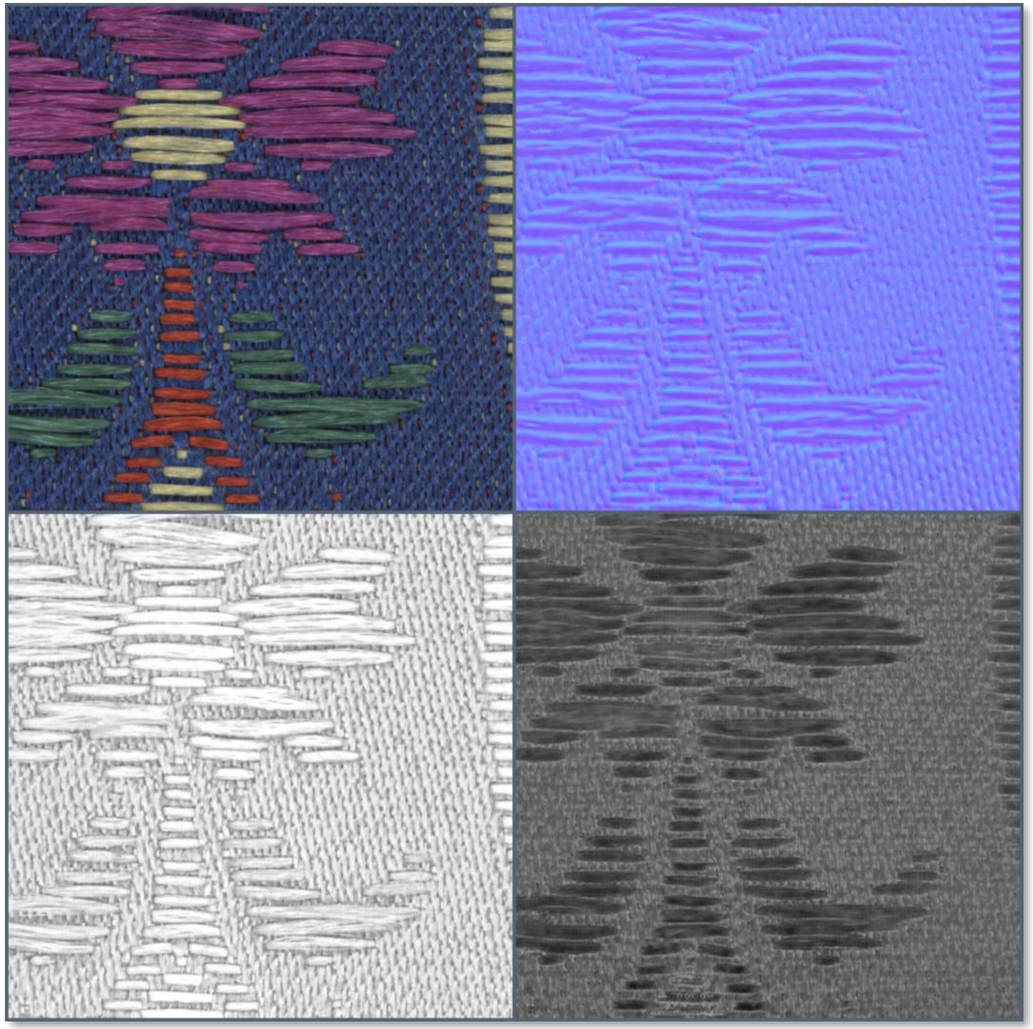
Uncertainty Quantification: Results



Active Learning: Results



Qualitative Results

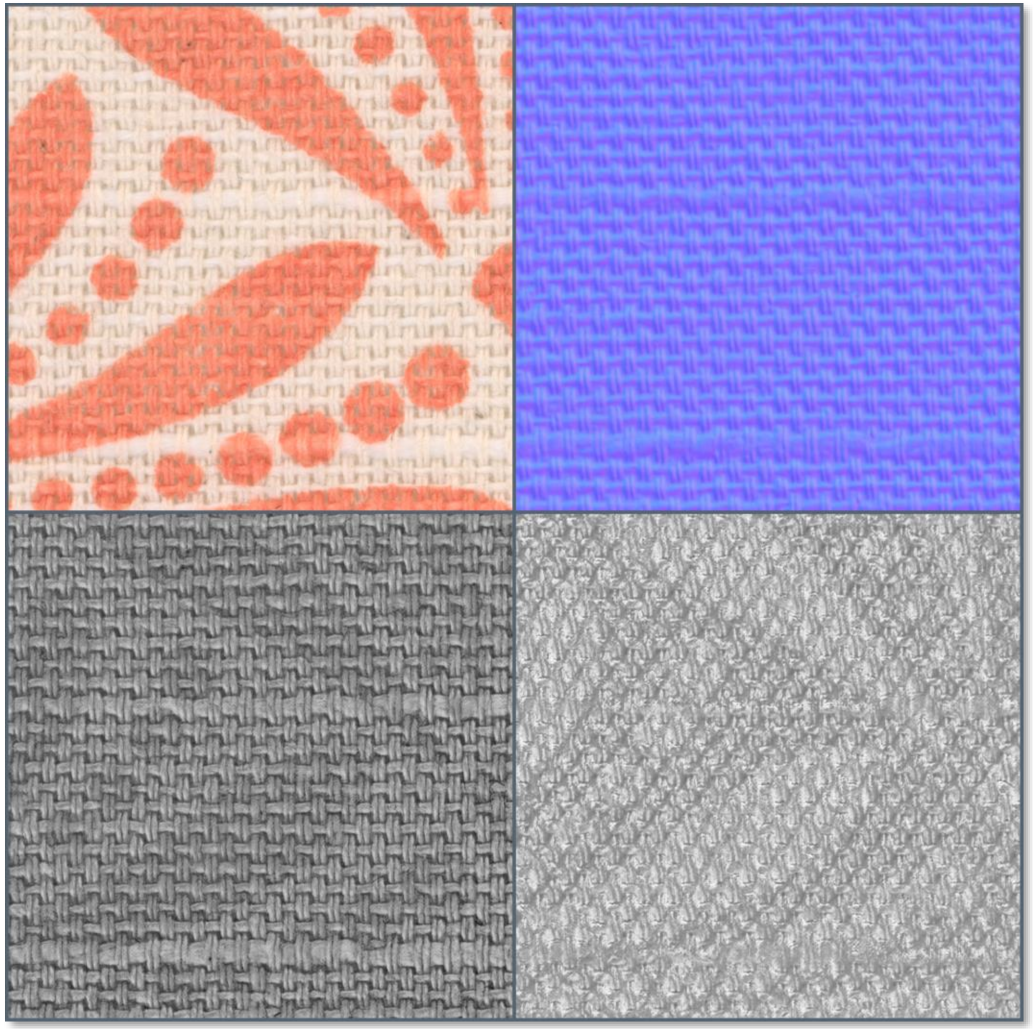


1000 PPI SVBRDF estimation (2x2 cm crop)

Render

Qualitative Results

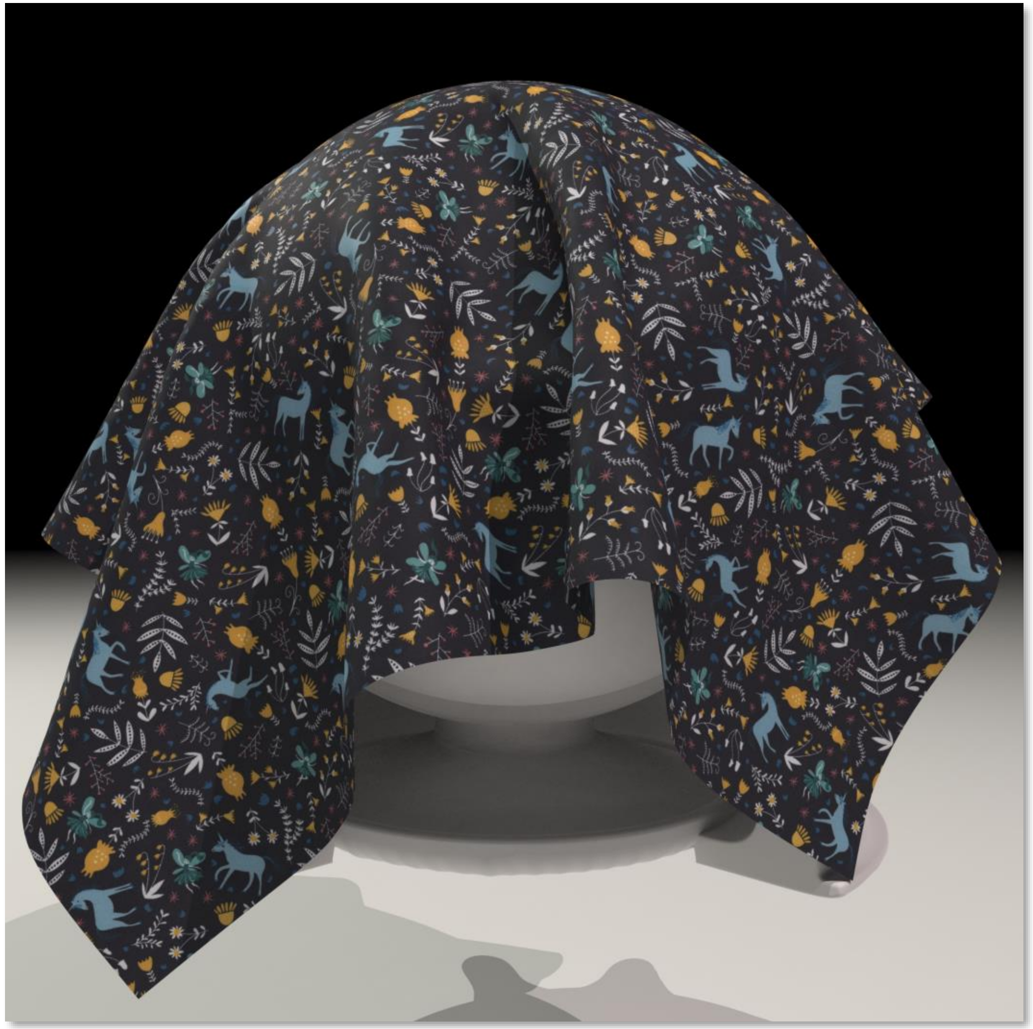
1000 PPI SVBRDF estimation (2x2 cm crop)



Render

Qualitative Results

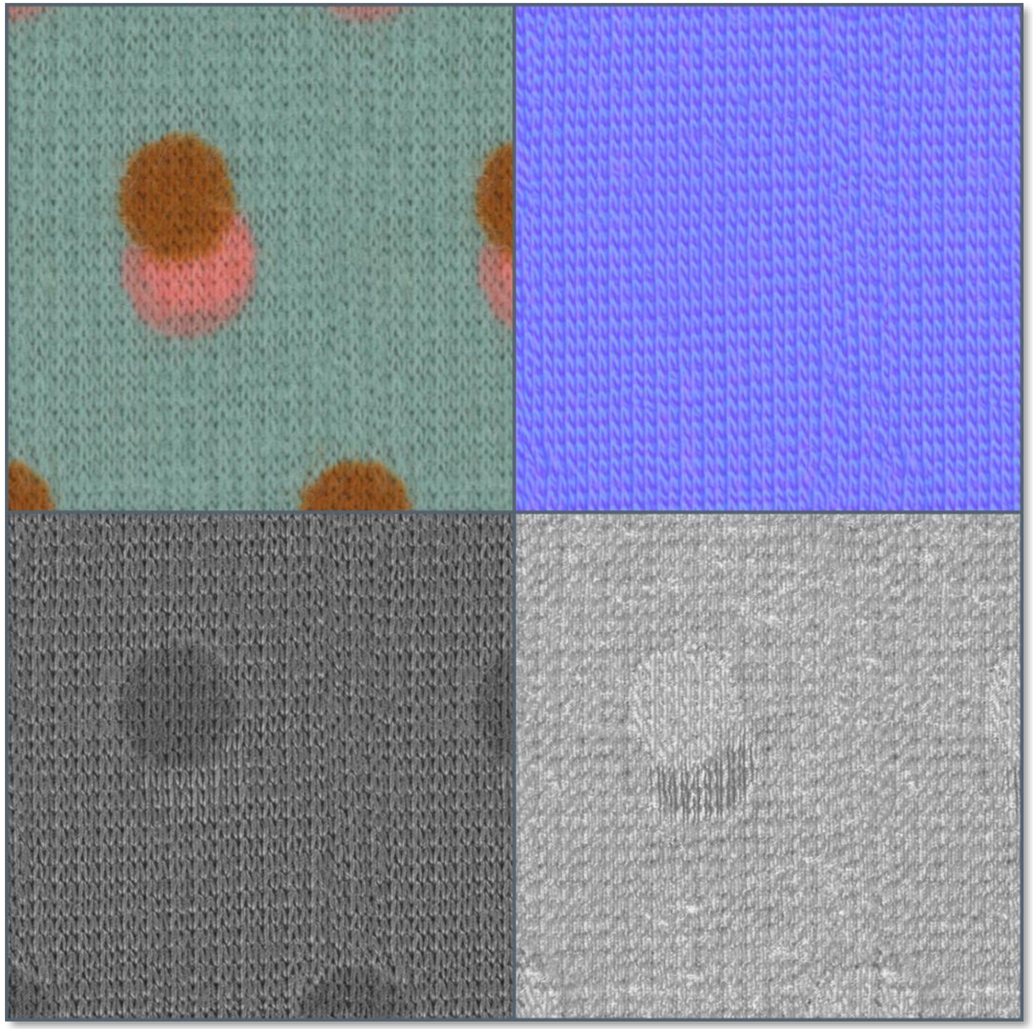
1000 PPI SVBRDF estimation (2x2 cm crop)



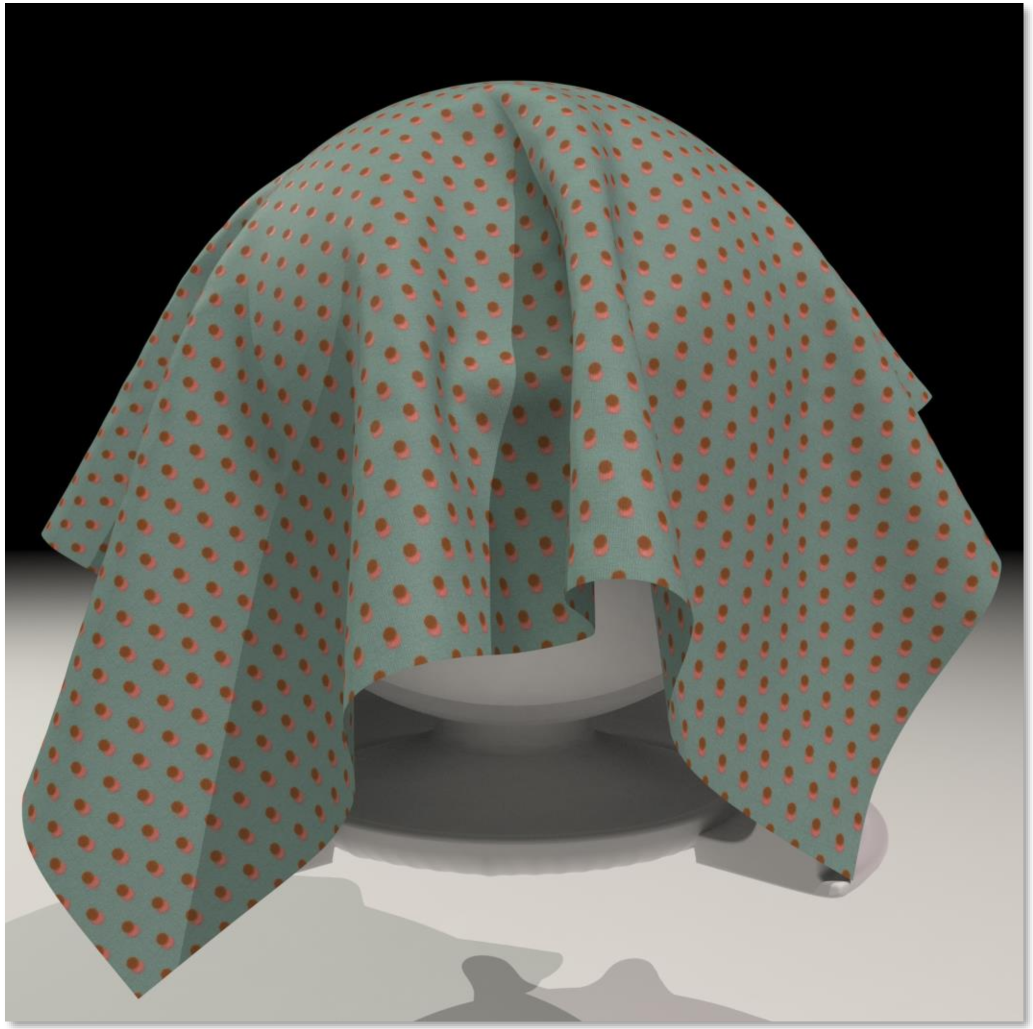
Render

Qualitative Results

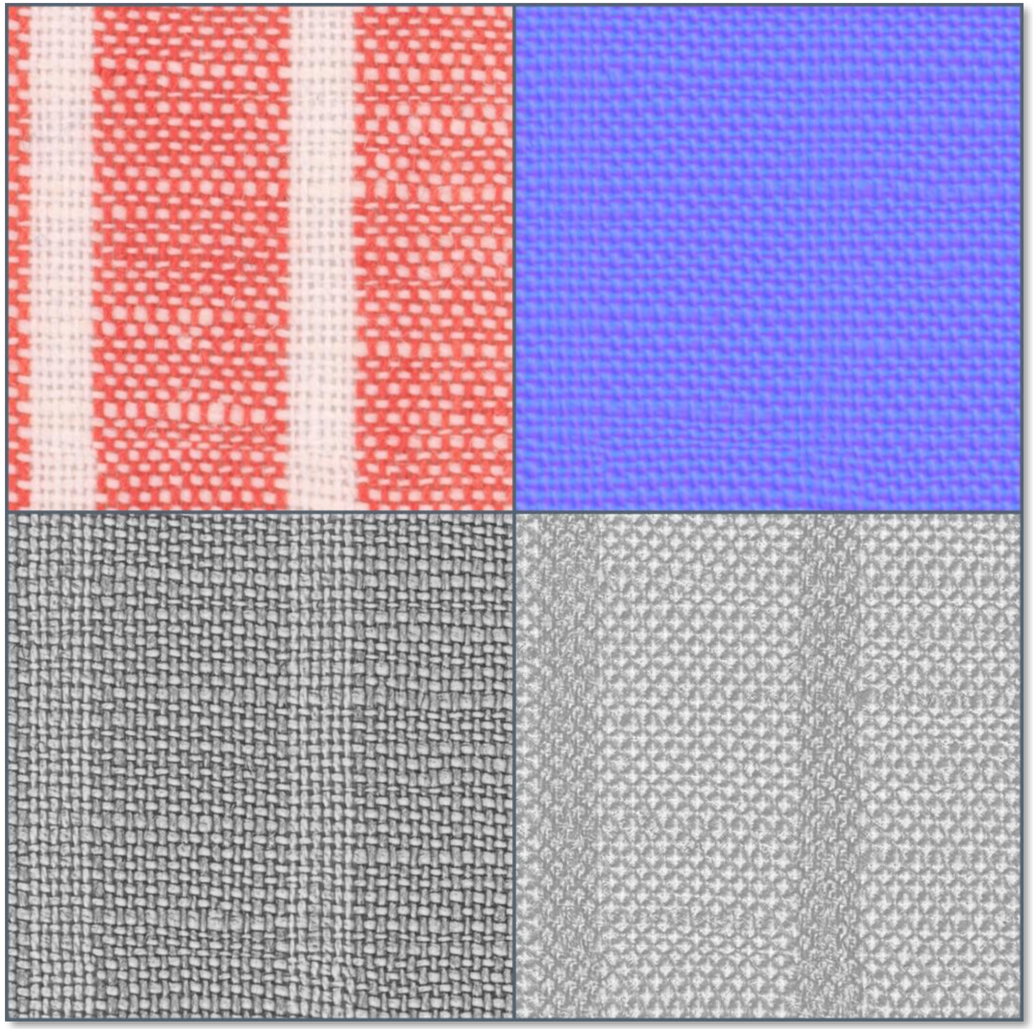
1000 PPI SVBRDF estimation (2x2 cm crop)



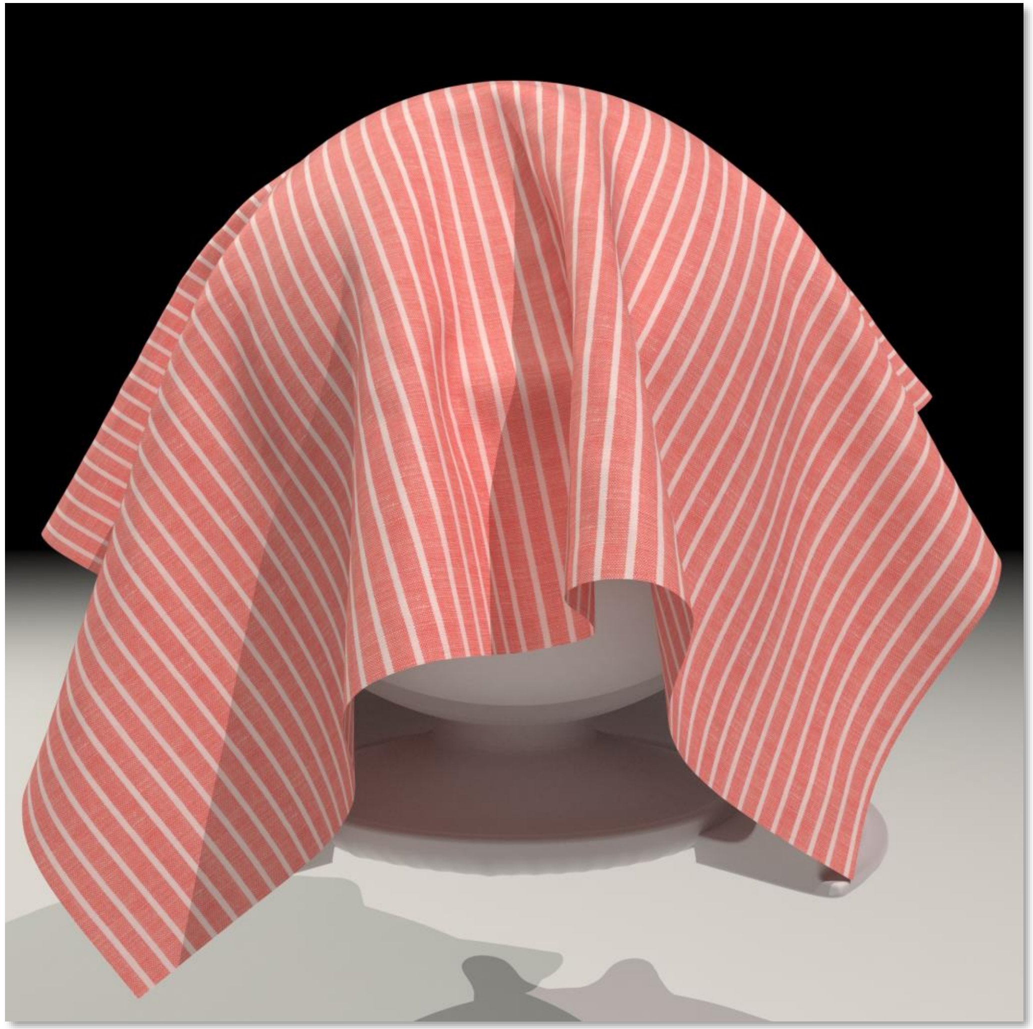
Render



Qualitative Results

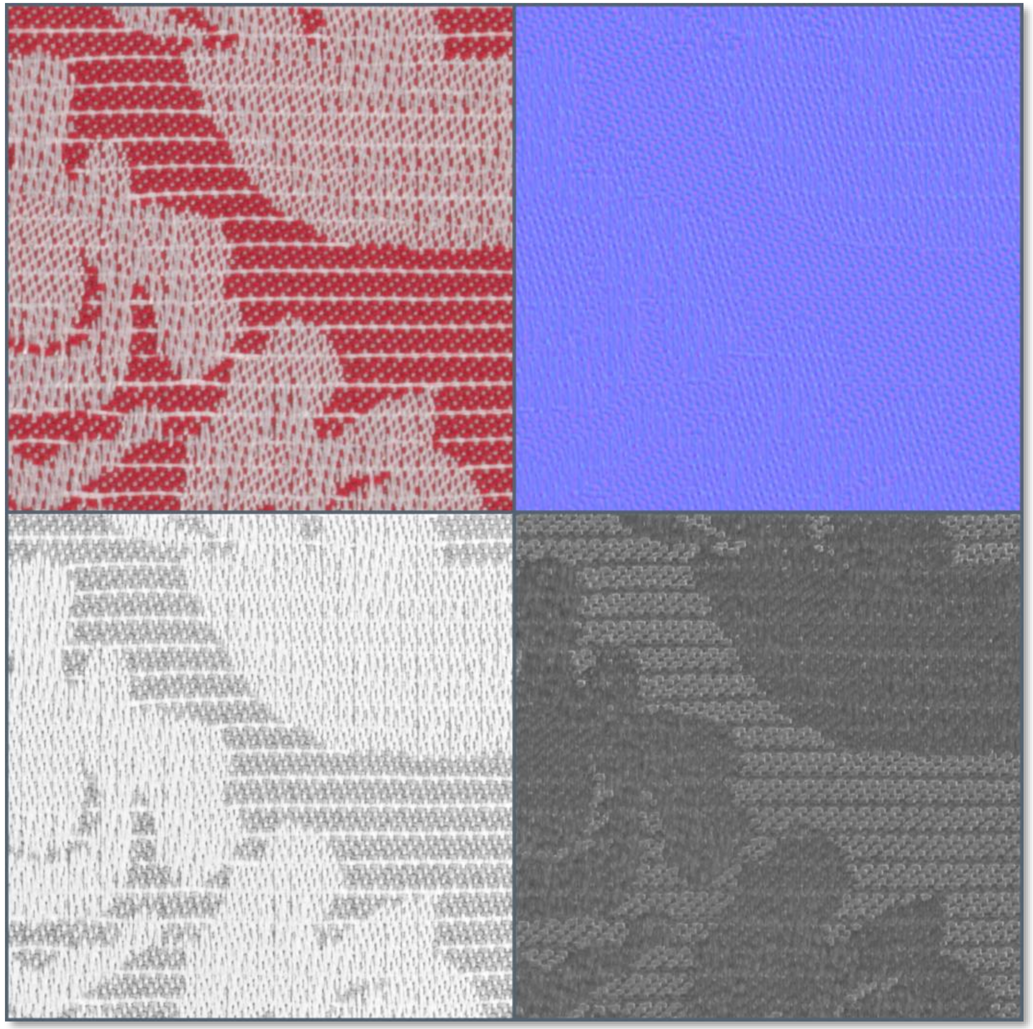


1000 PPI SVBRDF estimation (2x2 cm crop)

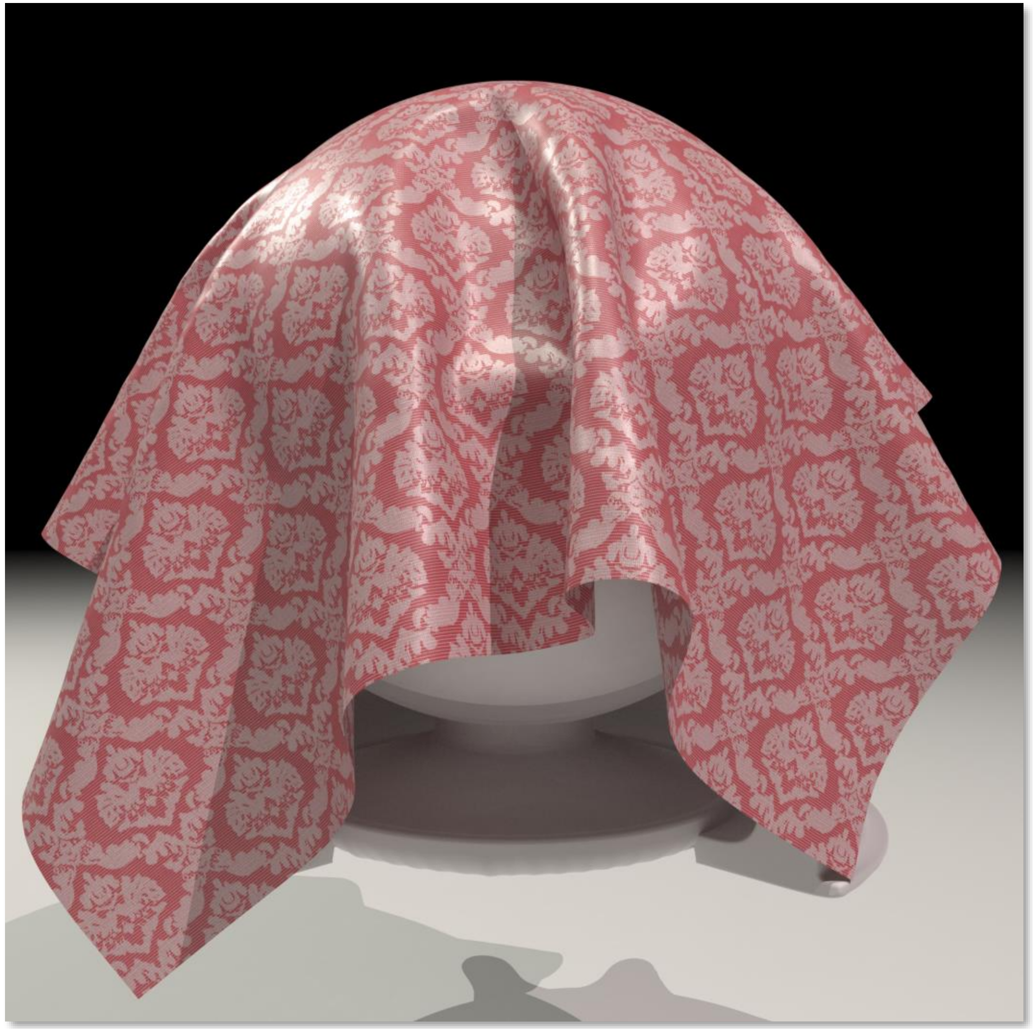


Render

Qualitative Results



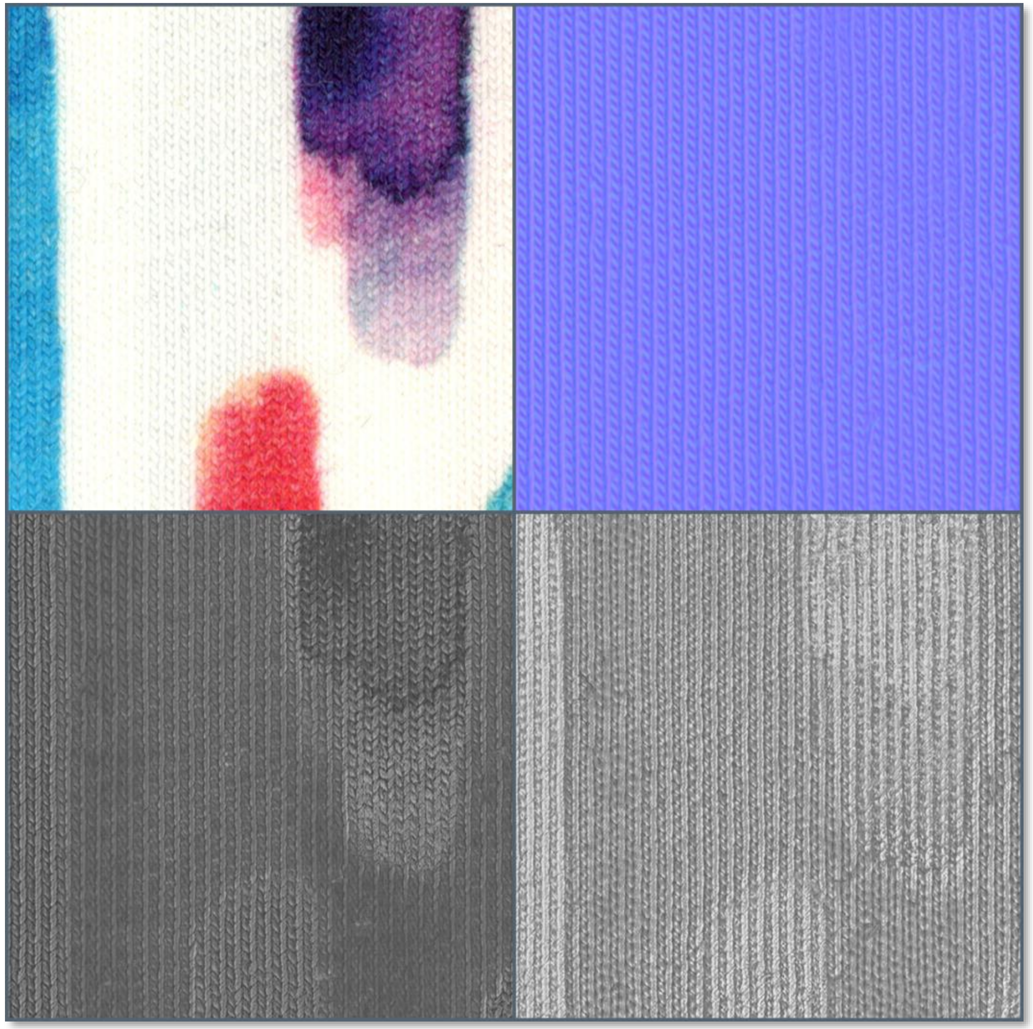
1000 PPI SVBRDF estimation (2x2 cm crop)



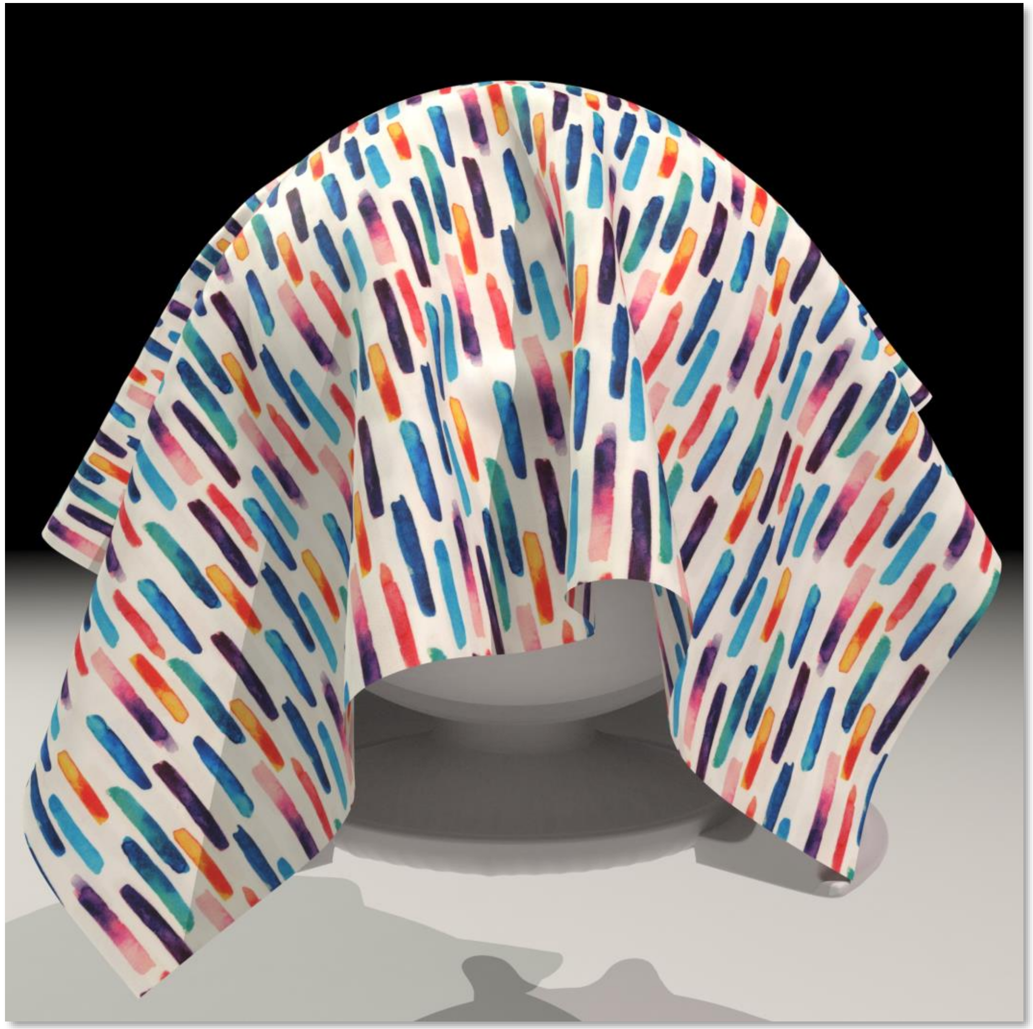
Render

Qualitative Results

1000 PPI SVBRDF estimation (2x2 cm crop)

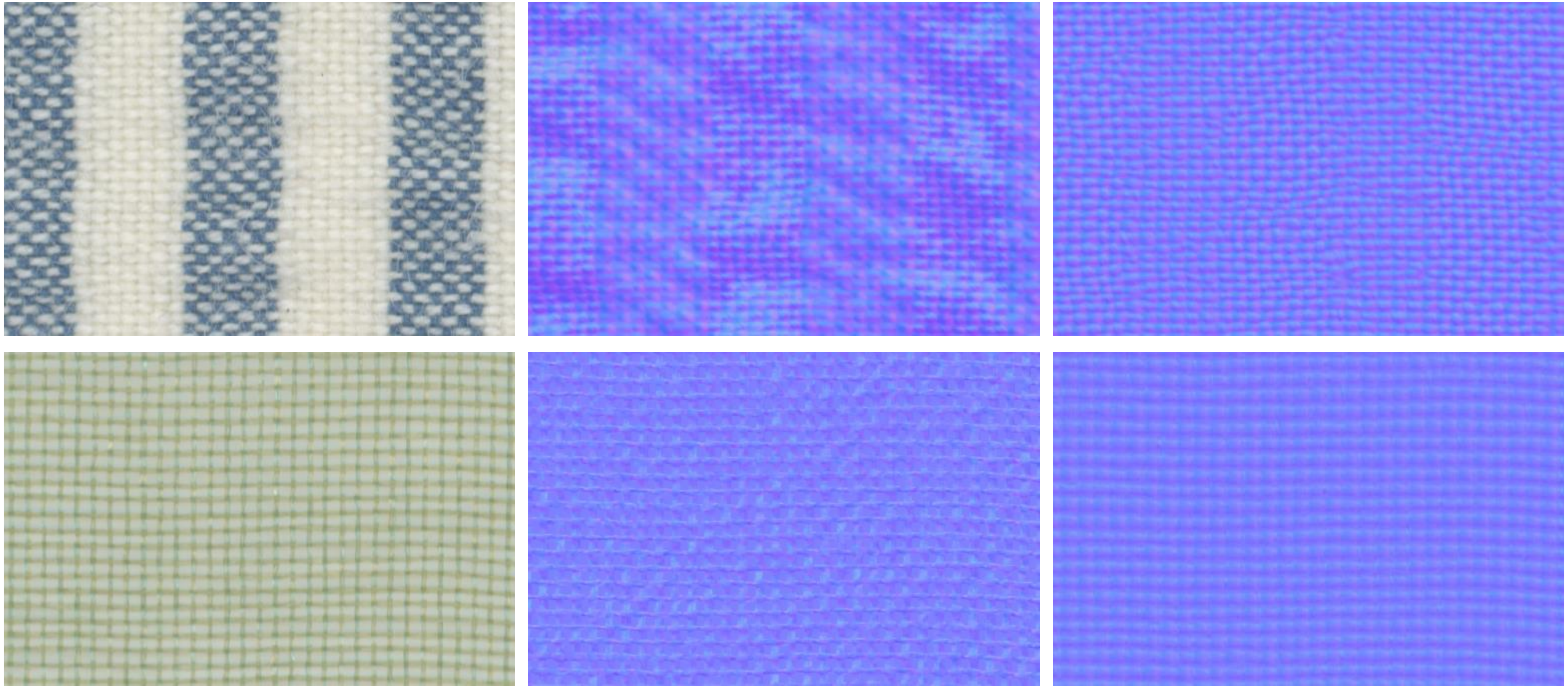


Render



Failure Cases

Input Image | Ground Truth | Estimation



Conclusions

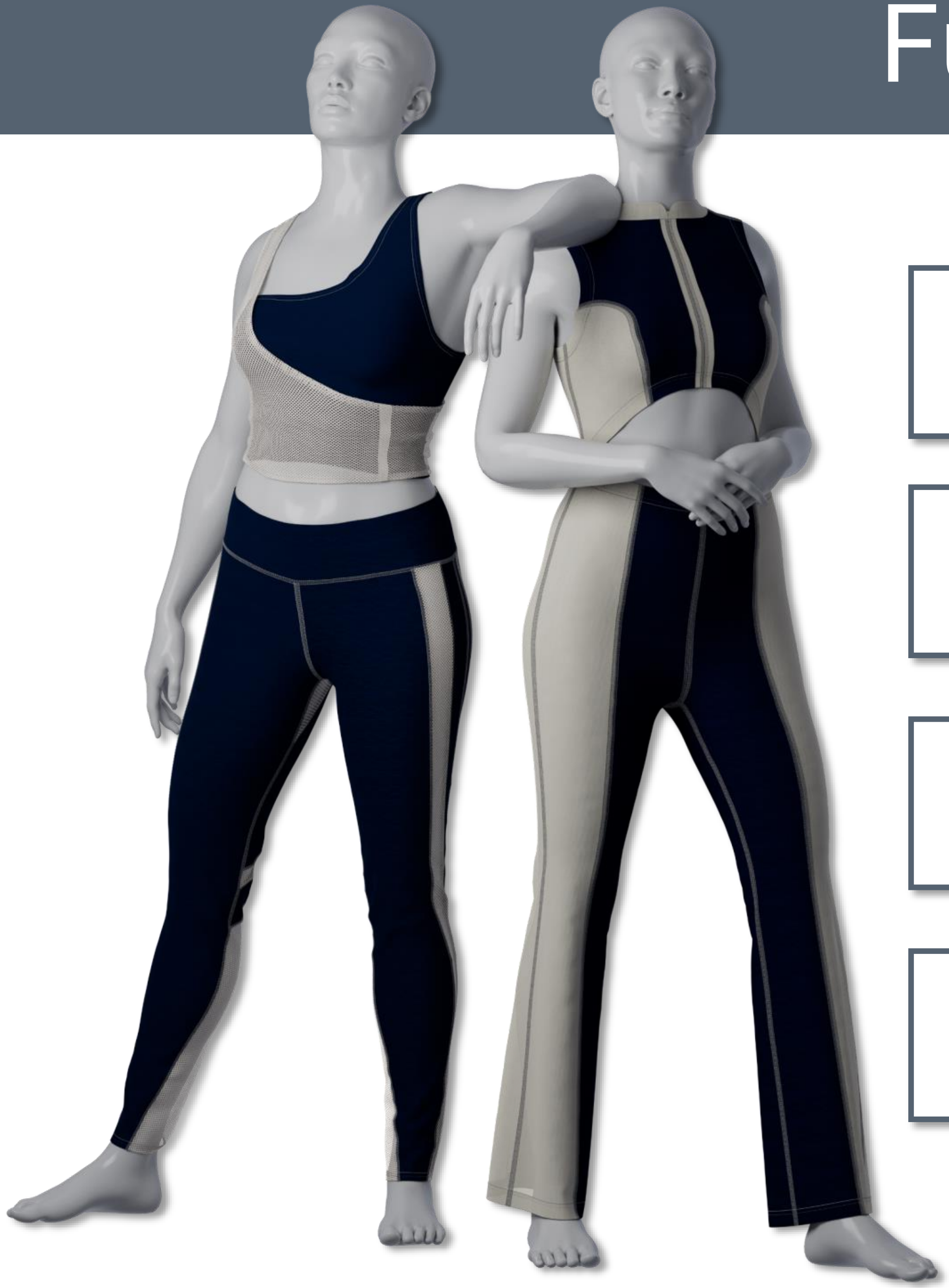
Generative model tailored for high resolution material digitization

Flatbed scanners provide very high resolution inputs

Microgeometry is a powerful cue for reflectance estimation

First uncertainty quantification method for SVBRDF estimation, increasing robustness and data effiency

Future Work



Estimate albedos

Expand material model: Transmittance, anisotropy

Increase dataset size and variety

Allow for other scanning devices

Additional Information

https://carlosrodriguezpardo.es/projects/UMat/

Thank you!

UMat: Uncertainty-Aware Single Image High Resolution Material Capture

CVPR 2023

Paper ID 7763

TUE-PM-156

SEDDI

Universidad Rey Juan Carlos