

# ROSA: Reconstructing Object Shape and Appearance Textures by Adaptive Detail Transfer – *Supplementary Material* –

Julian Kaltheuner    Patrick Stotko    Reinhard Klein  
University of Bonn  
{kaltheun, stotko, rk}@cs.uni-bonn.de

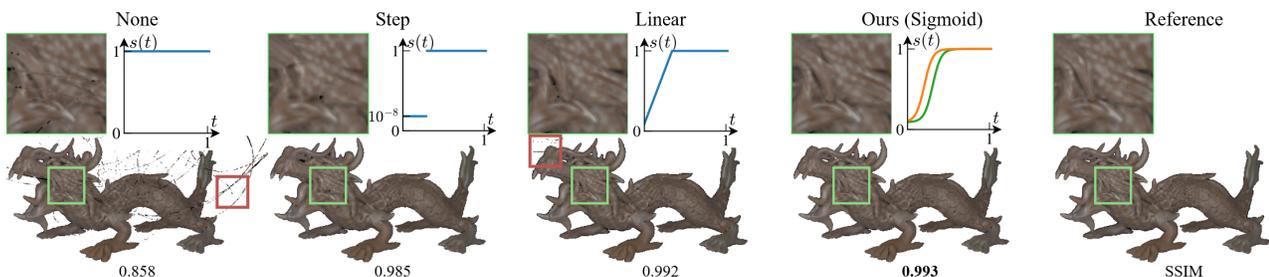


Figure 1. We show the behavior of our optimization with multiple different dampening functions. While our sigmoid scaling is able to reconstruct the object, other choices either lead to geometry artifacts or to a loss in detail.

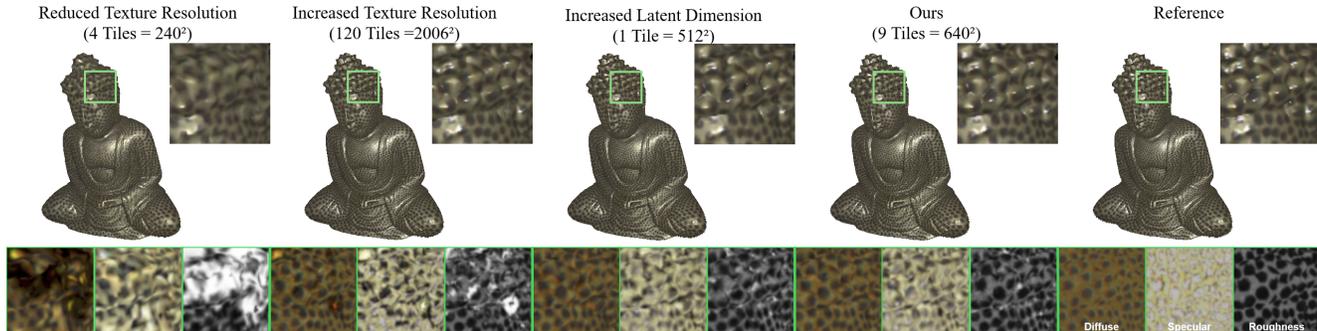


Figure 2. Comparison of estimated appearance feature with different amounts of texture tiles. Choosing a too high or low resolution results in artifacts whereas our approach achieves a similar quality compared to a single but manually tuned latent dimension.

## A. Optimization Stability

To keep the optimization procedure stable, we introduced the scaling of our losses  $s_{\text{loss}}$  and mesh adaption control variable  $s_{\text{control}}$ . In Fig. 1, we show the importance of the scaling term comparing the reconstructed surface geometry. Choosing a simple step function leads to less detailed reconstruction results. Changing the scaling function to a linear function or performing no dampening leads to deformations of the geometry that manifest into thread-like artifacts (highlighted in red). Meanwhile, our reconstruction stays stable and is able to produce a highly detailed geom-

etry reconstruction. We shift  $s_{\text{control}}$  (green) in time relative to  $s_{\text{loss}}$  (orange) to put more weight on the image loss which results in increased stability and dampens the geometry adaption in the early optimization phase.

## B. Texture Resolution

Our method is able to determine the appropriate texture size, since textures of large sizes can lead to hallucinated features while textures that are too small can lead to blurred features. In Fig. 2, we highlight the advantages of our reasonably computed total texture resolution compared to

other texture resolutions, where especially lower resolution lead to reconstruction artifacts in the estimated material and blurry re-renderings. Higher output resolutions, achieved by increasing the number of latent codes  $z$  to 120, produce sharp images that are consistent with the input images, but hallucinate high frequency features in the material. Due to the nature of the decoder architecture, we are also able to produce different resolution of output tiles, by changing the resolution of  $z$ . We highlight the ability of the decoder to produce high quality output materials with a single higher resolution tile and a comparable output resolution.