

# Multi-view Reconstruction of Highly Specular Surfaces in Uncontrolled Environments

## Supplementary Material

### A Probability distributions

**Multivariate t-distribution on  $\mathbb{R}^3$ .** A multivariate t-distribution is a generalization of the Student's t distribution to higher dimensions. Its probability density function (pdf) is

$$t(\mathbf{x}|\boldsymbol{\mu}, \nu, \boldsymbol{\Sigma}) = \frac{\Gamma[(\nu + p)/2]}{\Gamma(\nu/2)\nu^{p/2}\pi^{p/2} |\boldsymbol{\Sigma}|^{1/2} \left[1 + \frac{1}{\nu}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right]^{(\nu+p)/2}}, \quad (1)$$

where the  $p = 3$  is the dimensionality of the domain,  $\nu \in \mathbb{R}$  is the degrees of freedom,  $\boldsymbol{\mu} \in \mathbb{R}^3$  is the mean vector and  $\boldsymbol{\Sigma}$  is the correlation matrix [1, Chapter 1.1]. We use a diagonal correlation matrix,  $\boldsymbol{\Sigma} = \sigma^2 I_3$ , and write the pdf as

$$t(\mathbf{x}|\boldsymbol{\mu}, \nu, \sigma^2) = \frac{\Gamma[(\nu + 3)/2]}{\Gamma(\nu/2)\nu^{3/2}\pi^{3/2}\sigma^3 \left[1 + \frac{1}{\sigma^2\nu}\|\mathbf{x} - \boldsymbol{\mu}\|^2\right]^{(\nu+3)/2}}. \quad (2)$$

We used  $\nu = 0.01$  and  $\sigma = 0.01$  throughout all our experiments.

**von Mises-Fisher distribution on  $S^2$ .** The 2 dimensional von Mises-Fisher distribution (also called the Fisher distribution) is a probability distribution over the unit sphere  $S^2 \subset \mathbb{R}^3$ . Its pdf is

$$v(\mathbf{x}|\boldsymbol{\mu}, \kappa) = \frac{\kappa}{4\pi \sinh \kappa} e^{\kappa \boldsymbol{\mu}^T \mathbf{x}} = \frac{\kappa}{2\pi(e^\kappa - e^{-\kappa})} e^{\kappa \boldsymbol{\mu}^T \mathbf{x}}, \quad (3)$$

where  $\boldsymbol{\mu} \in \mathbb{R}^3$  is the mean direction and  $\kappa \in \mathbb{R}$  is the concentration parameter [2, Chapter 9.3]. As  $\mathbf{x}, \boldsymbol{\mu} \in \mathbb{R}^3$ , we can write  $\boldsymbol{\mu}^T \mathbf{x} = \cos \theta$ , where  $\theta$  is the angle between  $\mathbf{x}$  and  $\boldsymbol{\mu}$ . Thus,

$$v(\theta, \kappa) = \frac{\kappa}{2\pi(1 - e^{-\kappa})} e^{\kappa(\cos \theta - 1)} \quad (4)$$

Furthermore, when  $\kappa$  is large,  $e^{\kappa x}$  converges to zero very quickly and for small  $\theta$  we can write  $\cos \theta = 1 - \frac{\theta^2}{2} + O(\theta^2)$ . As a consequence,

$$v(\theta, \kappa) \approx \frac{\kappa}{2\pi} e^{-\kappa \frac{\theta^2}{2}}. \quad (5)$$

Thus, for small angles  $\theta$  and large concentration  $\kappa$  a von Mises-Fisher distribution can be seen as a scaled normal distribution with its mean at zero and the variance  $\sigma^2 = 1/\kappa$ .

We use this observation in Section 3.2.1 to set our parameter  $\kappa = \theta_d^{-2}$  where  $\theta_d$  is roughly one standard deviation of a scaled normal distribution over the angles. Thus our remodeling of the continuous density in log-space can be seen as a Gaussian blur on  $S^2$  with a variance  $\theta_d^2 = 1/\kappa$  where we initialize  $\theta_d$  to  $5^\circ$  and linearly decrease it to reach  $2^\circ$ , its minimum value, at the fifth iteration of the entire optimization process.

## B Comparison with Oxholm and Nishino [3]

As stated in Section 4.2 of the main body of the paper, we cannot directly compare our methods as Oxholm and Nishino [3] were unable to run their method on new datasets and it is not possible to run our method on their dataset due to poorly calibrated images. To compare our methods as faithfully as possible, we implemented three variants of their algorithm. Instead of directly using the probability distributions over the surface normals to drive the shape of the object, all variants use our two-step optimization. That is, we first extract representative normals for each vertex on the mesh and then refine the vertex locations to better explain the normals.

**Method 1** This is the method we compare against in the main body of the paper. The pipeline is identical to ours but we use their method to compute the distributions over the normals, which does not account for inter-reflections. More precisely, we compute the per-vertex distributions over the possible surface normals using the method described in [3, Section 3.1], but use our own smoothness term instead of their priors  $p_c$ ,  $p_a$  and  $p_e$ . We also lower the sampling rate of normal space to match theirs by setting the resolution exponent to  $N_r = 4$ . This results in  $N_p = 3072$  samples which corresponds to roughly twice the number of samples used in [3].

**Method 2** Their method is used to compute the distributions over the normals, as in Method 1, and their area priors  $p_a$  and  $p_e$  are used when refining the surface instead of using our volume and frontier point terms. The vertices are allowed to be displaced *only along their representative normal* during the optimization. We do not perform our post-processing step after optimizing the surface of the object, instead we recenter and rescale the mesh as described in [3, Section 4].

**Method 3** This method is closest to the one described in [3], it is similar to method 2 but the displacement of each vertex *is not constrained* during the optimization.

As can be seen in Figures 3, 4 and 5 the first method produces noisy results and is unable to carve deep concavities, such as the palm of HEAVYMETAL in Figure 4. Method 2 is much noisier and diverges quickly from an acceptable solution while Method 3 is heavily biased and makes the mesh degenerate after more than two iterations.

## C Influence of silhouette quality on the reconstruction error

Our algorithm generates the initial surface from and computes the frontier points using user-made silhouettes. To gage the impact of silhouette quality on the reconstruction, we performed two experiments where we perturb the silhouettes of the synthetic object BUNNY in the SPONZA scene. One where the perturbations are constrained to be conservative, i.e. they can only grow the silhouette. The other one performs unconstrained perturbations on the silhouette. The perturbations are performed by adding Gaussian noise to the contour points with a standard deviation of  $\sigma$  pixels. See Figure 1 for example silhouettes using different noise levels.

Figure 2 shows, conservative perturbations are easier to recover from. For noise levels of 2 pixels or less, we even converge to result as the ground truth silhouette. On the other hand, unconstrained noise it harder to recover from, partly because the initial guess produced by space carving is smaller than the ground truth mesh.

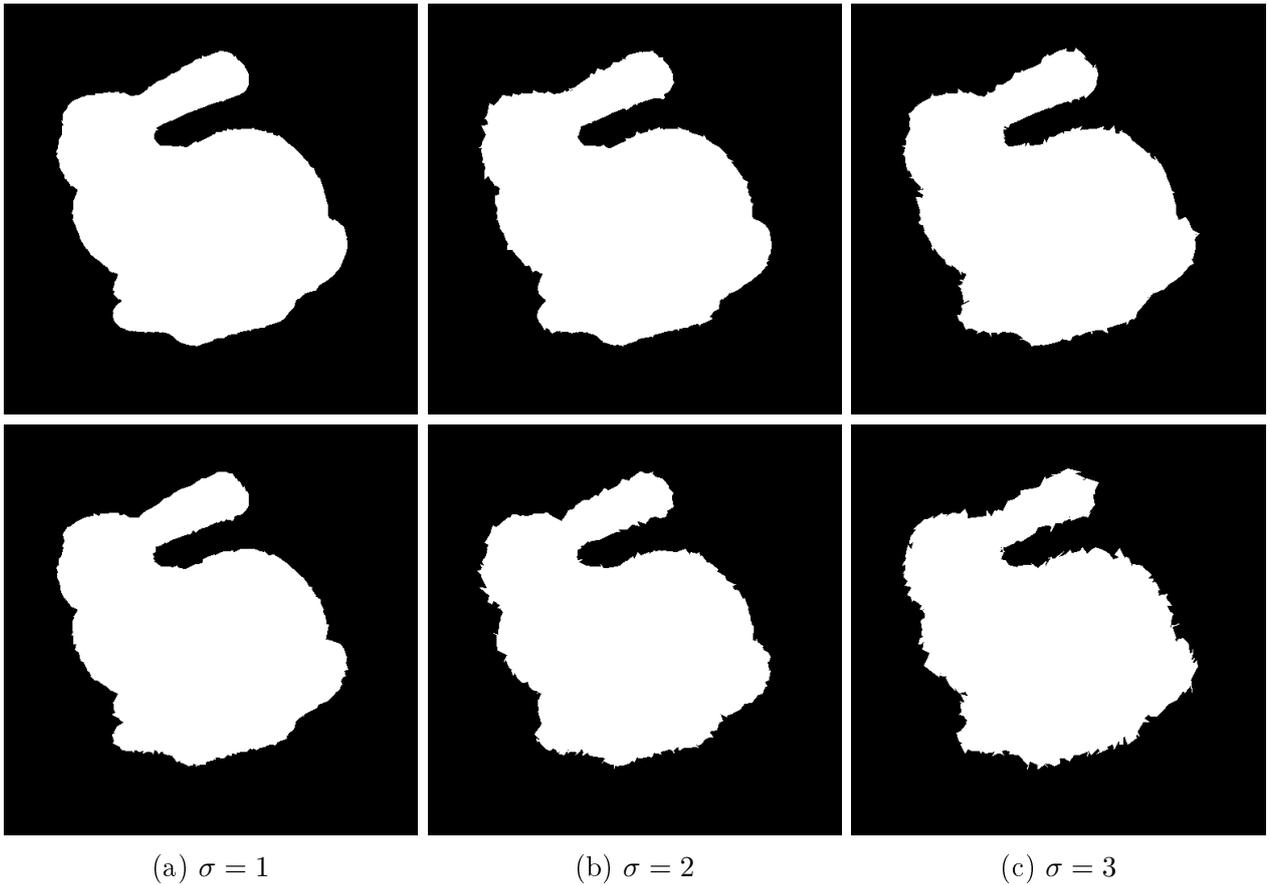


Figure 1: Example of corrupted silhouettes using different levels of noise. The silhouettes in the upper row were conservatively, while the silhouettes in the lower row were modified using unconstrained perturbations.

## References

- [1] S. Kotz and S. Nadarajah. *Multivariate  $t$ -distributions and their applications*. Cambridge University Press, 2004.
- [2] K. V. Mardia and P. E. Jupp. *Directional Statistics*. John Wiley & Sons, Inc., 2008.
- [3] G. Oxholm and K. Nishino. Multiview shape and reflectance from natural illumination. In *CVPR*, pages 2163–2170, June 2014.

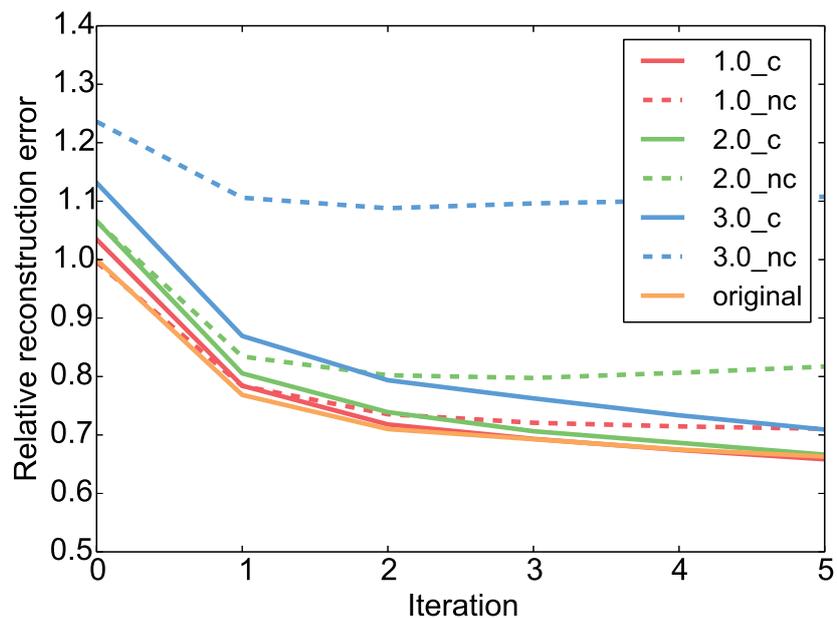


Figure 2: The relative reconstruction error with respect to the space carving performed using accurate silhouettes for BUNNY in SPONZA. The number in the labels correspond to  $\sigma$  of the perturbations, C corresponds to conservative perturbations and NC corresponds to unconstrained perturbations.

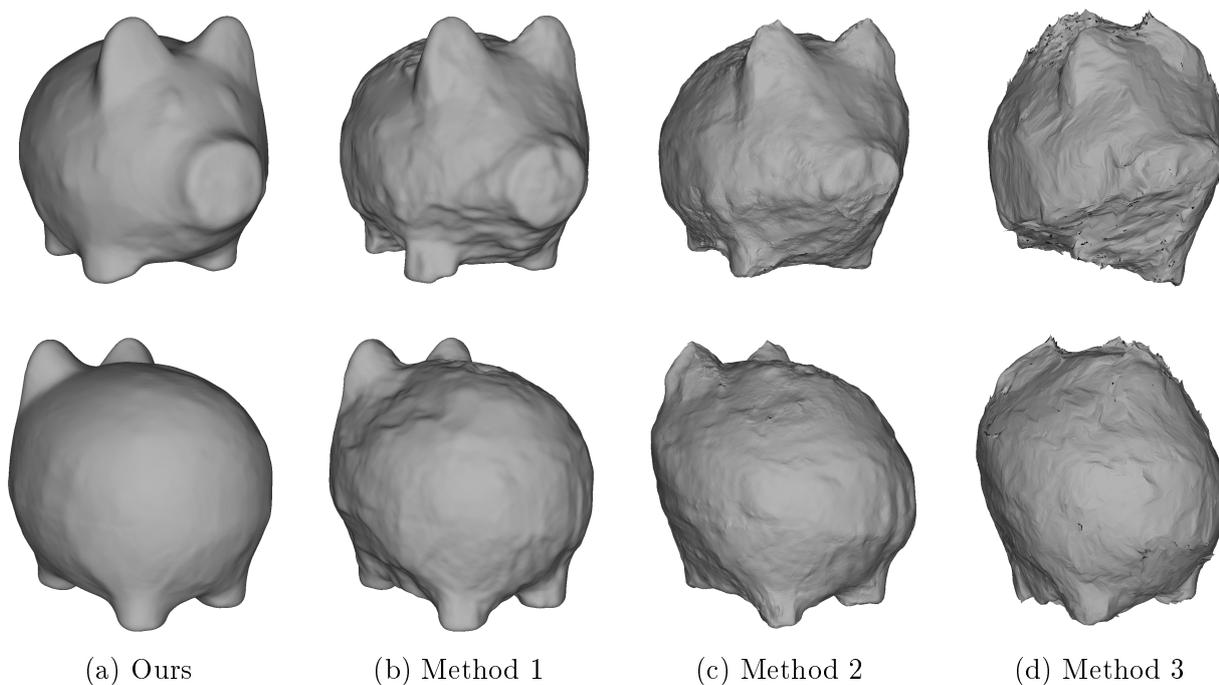


Figure 3: Results for PIGGY in THEATER after 5 iterations for (a,b,c) and 2 iterations for (d)

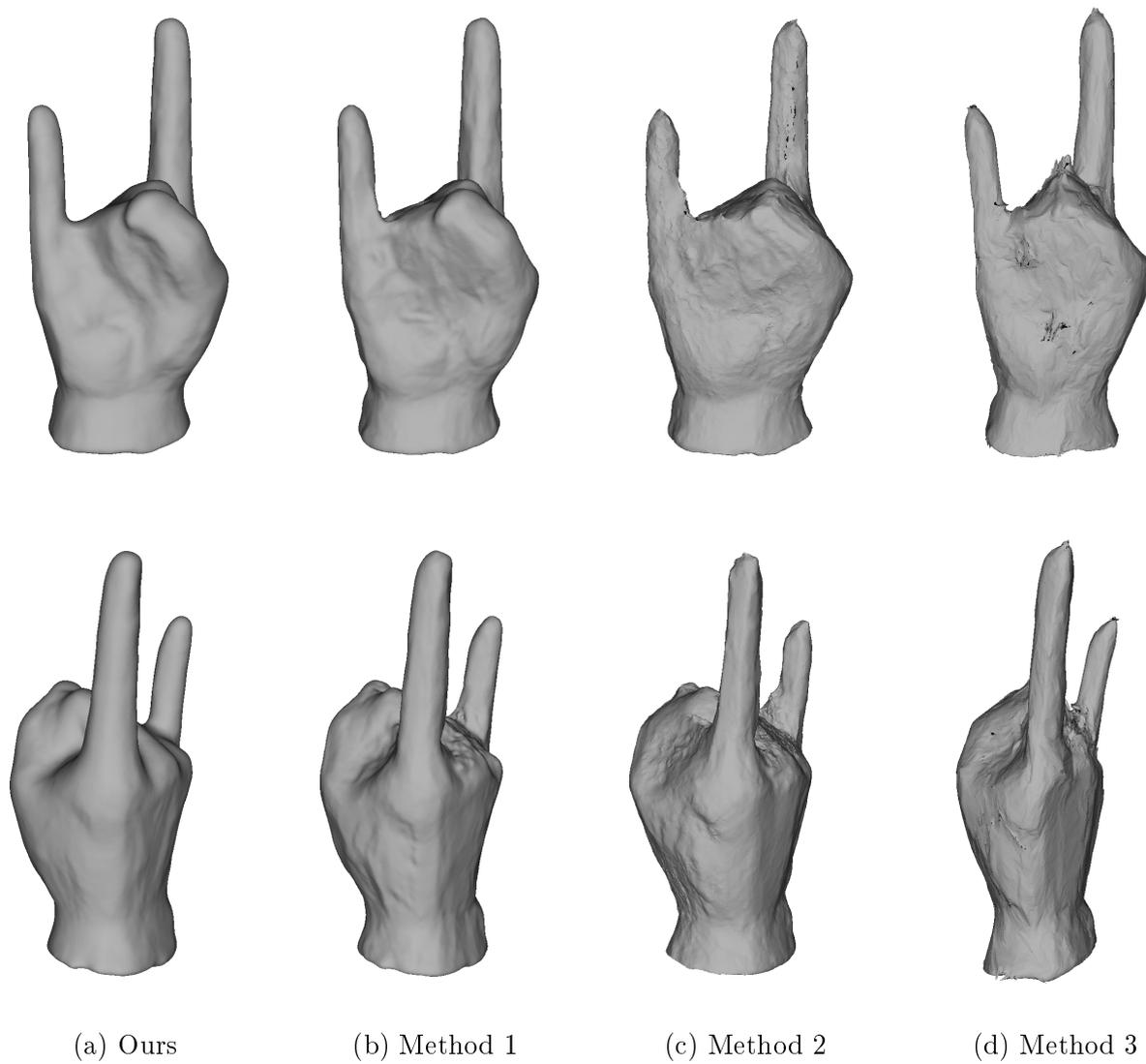


Figure 4: Results for HEAVYMETAL in THEATER after 5 iterations for (a,b,c) and 2 iterations for (d)

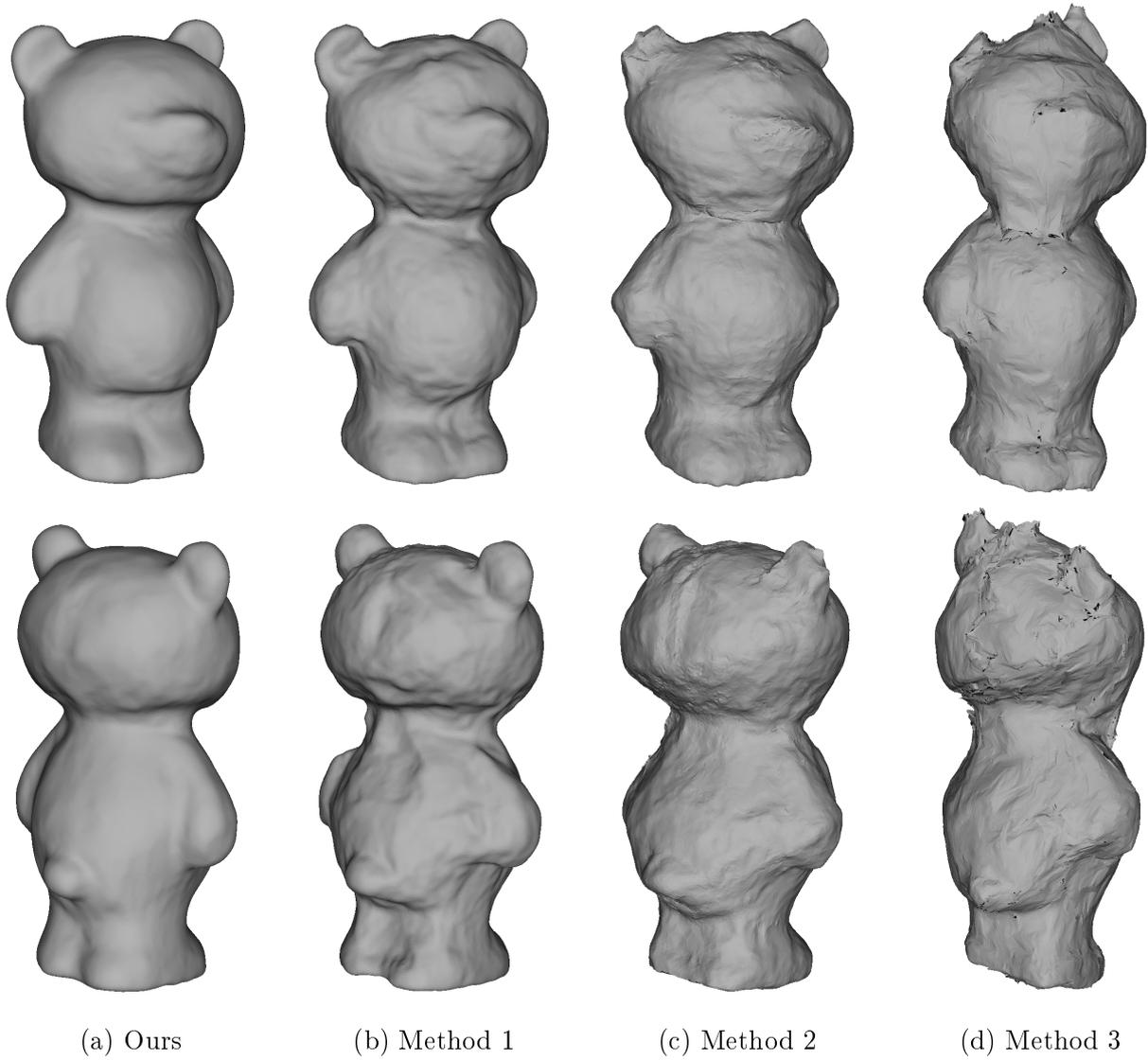


Figure 5: Results for TEDDY in THEATER after 5 iterations for (a,b,c) and 2 iterations for (d)