An Adaptive Screen-Space Meshing Approach for Normal Integration

Supplementary Material

Moritz Heep and Eduard Zell

University of Bonn, Bonn, Germany

Overview

The goal of this supplementary material is to provide additional technical details of the proposed meshing and integration pipeline and to present further evaluation results. To simplify navigation, we follow the structure of the manuscript.

The implementation can be found under moritzheep.github.io/adaptive-screen-meshing.

Table of Contents

1 Fundamental Forms in the Perspective Case

In the perspective case, the surface tangents are given by

$$\partial_u \boldsymbol{x} = \left(\partial_u \boldsymbol{r} - \frac{\boldsymbol{n} \cdot \partial_u \boldsymbol{r}}{\boldsymbol{n} \cdot \boldsymbol{r}} \cdot \boldsymbol{r} \right) \cdot z \,, \qquad \partial_v \boldsymbol{x} = \left(\partial_v \boldsymbol{r} - \frac{\boldsymbol{n} \cdot \partial_v \boldsymbol{r}}{\boldsymbol{n} \cdot \boldsymbol{r}} \cdot \boldsymbol{r} \right) \cdot z \,, \tag{1}$$

i.e. they scale linearly with depth. As the fundamental forms are given by

$$\boldsymbol{I}_{ij} = \partial_i \boldsymbol{x} \cdot \partial_j \boldsymbol{x} \,, \qquad \boldsymbol{II}_{ij} = -\partial_i \boldsymbol{x} \cdot \partial_j \boldsymbol{n} \,, \tag{2}$$

we get $\boldsymbol{I} \propto z^2$ and $\boldsymbol{II} \propto z$. From the scaling behaviour of the first fundamental form, it follows that all lengths we measure on screen are proportional to the depth z. This is simply the well-known relationship that the screen-size of an object is inversely proportional to the distance to the camera.

For the generalized eigenvalue problem

$$\kappa_i \cdot \boldsymbol{I}\, \boldsymbol{v}_i = \boldsymbol{II}\, \boldsymbol{v}_i \,, \tag{3}$$

this means that the two eigenvalues κ_i are proportional to z^{-1}.

Finally, if we look at the error estimate

$$L = \sqrt{\frac{6\epsilon}{|\kappa|} - \epsilon^2} \tag{4}$$

we can choose the approximation $\epsilon = \lambda \cdot z$ in relation to depth. Then, the optimal length L will be proportional to z. Since the actual lengths we measure on screen, are also proportional to z, the ratio of optimal length and actual length is *not* dependant on z. In practice, we can set $z = 1$ and treat ϵ as the relative error compared to z.

2 Mesh-Based Integration - Full Derivation

In the unified integration framework, the perspective case reads

$$E_{\text{Int}} = \int_\Omega (\boldsymbol{n} \cdot \boldsymbol{r}\, \partial_u z + n_x)^2 + (\boldsymbol{n} \cdot \boldsymbol{r}\, \partial_v z + n_y)^2 \; du\, dv \,. \tag{5}$$

where $\boldsymbol{r}$ is the ray given by the camera matrix and z is logarithmic depth. In mesh-based integration, we want to obtain depth values $z_1, ..., z_{|\mathcal{V}|}$ for each vertex. Depth for any point on a given triangle is defined by linear interpolation using barycentric coordinates. We start by rewriting the integrand in vector form:

$$E_{\text{Int}} = \int_\Omega \left\| \boldsymbol{n} \cdot \boldsymbol{r}\, \nabla z + \nabla \boldsymbol{u}^t \cdot \begin{pmatrix} n_x \\ n_y \end{pmatrix} \right\|^2 \; du\, dv \,. \tag{6}$$

As we use linear interpolation, the gradients are constant within each face. From this, it follows that

$$\int_f \nabla g \cdot \nabla h \, d\Omega = \frac{1}{2} \sum_{(i,j)\in\mathcal{E}_f} \cot(\alpha_k)(g_j - g_i)(h_j - h_i) \tag{7}$$

where the integral runs over the area of the face and α_k is the angle opposite to edge ij. This is a generalization of [8, Lemma 5]. Applying the polynomial theorem to the integration functional, we can discretize each addend using Eq. (7).

Quadratic Term

We start with the term that is quadratic in ∇z. Applying Eq. (7) gives

$$\int_f (\boldsymbol{n} \cdot \boldsymbol{r})^2 \cdot \|\nabla z\|^2 \, d\Omega = \left(\frac{1}{A_f} \int_f (\boldsymbol{n} \cdot \boldsymbol{r})^2 \, d\Omega \right) \sum_{(i,j)\in\mathcal{E}_f} \frac{\cot(\alpha_k)}{2} \cdot (z_j - z_i)^2 \,. \tag{8}$$

Assuming constant face normals $\boldsymbol{n}_f$, we can perform the remaining integral using [9, Theorem 2.2] to get

$$2m_f := \frac{1}{A_f} \int_f (\boldsymbol{n} \cdot \boldsymbol{r})^2 \, d\Omega = \frac{1}{6} \sum_{\substack{i,j\in\mathcal{V}_f \\ i\leq j}} (\boldsymbol{n}_f \cdot \boldsymbol{r}_i)(\boldsymbol{n}_f \cdot \boldsymbol{r}_j) \tag{9}$$

Linear Term

Applying the same strategy to the linear term, we get

$$\int_f (\boldsymbol{n} \cdot \boldsymbol{r}) \cdot \left(\begin{pmatrix} n_x \\ n_y \end{pmatrix} \cdot \nabla \boldsymbol{u} \right) \cdot \nabla z \, d\Omega \tag{10}$$

$$= \left(\frac{1}{A_f} \int_f \boldsymbol{n} \cdot \boldsymbol{r} \, d\Omega \right) \sum_{(i,j)\in\mathcal{E}_f} \frac{\cot(\alpha_k)}{2} \cdot (z_j - z_i) \begin{pmatrix} n_{f,x} \\ n_{f,y} \end{pmatrix} \cdot (\boldsymbol{u}_j - \boldsymbol{u}_i) \,. \tag{11}$$

Again, the remaining integral can be evaluated to be

$$b_f := \frac{1}{A_f} \int_f \boldsymbol{n} \cdot \boldsymbol{r} \, d\Omega = \frac{1}{3} \sum_{i\in\mathcal{V}_f} \boldsymbol{n}_f \cdot \boldsymbol{r}_i$$

Constant Term

Finally, we get

$$\int_f \left\| \nabla \boldsymbol{u}^t \cdot \begin{pmatrix} n_x \\ n_y \end{pmatrix} \right\|^2 du \, dv = \sum_{(i,j)\in\mathcal{E}_f} \frac{\cot(\alpha_k)}{2} \cdot \left(\begin{pmatrix} n_{f,x} \\ n_{f,y} \end{pmatrix} \cdot (\boldsymbol{u}_j - \boldsymbol{u}_i) \right)^2 \tag{12}$$

for the constant term.

Discretized Integration Energy

Putting all three terms together, each triangle contributes

$$\sum_{(i,j)\in\mathcal{E}_f} \cot(\alpha_k)\left(m_f\cdot\delta z_{ij}^2 + b_f\cdot\begin{pmatrix}n_{f,x}\\ n_{f,y}\end{pmatrix}\cdot\boldsymbol{\delta u}_{ij}\cdot\delta z_{ij} + \frac{1}{2}\left(\begin{pmatrix}n_{f,x}\\ n_{f,y}\end{pmatrix}\cdot\boldsymbol{\delta u}_{ij}\right)^2\right) \tag{13}$$

to the integration energy where we use the short hand notation $\delta x_{ij} = x_j - x_i$. Hence, the total integration energy is the sum over all triangle contributions.

To obtain the optimality condition reported in the main work, we need to take the derivative with regard to a single z_k. As the discretized integration energy E_{Int} only contains differences between adjacent vertices, $\frac{\partial E_{\mathrm{Int}}}{\partial z_k}$ may only depend on z_k and its neighbours $z_i, i \in \mathcal{V}_k$. Collecting the terms for each neighbour, we arrive at the condition that was reported in the main work.

In the perspective case, solving the linear system gives one logarithmic depth value per vertex and we get the actual depth by taking the exponential. In the orthographic case, we replace $\boldsymbol{r} \to \boldsymbol{e}_3$ and the solution of the linear system is depth, no exponentiation needed.

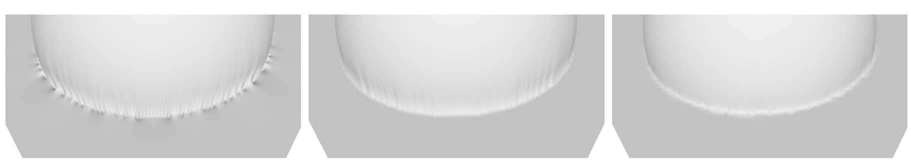
3 Gibbs Phenomenon

To further investigate the actual cause of the Gibbs phenomenon in normal integration, we analyze two methods [4, 5]. For simplicity, we focus on the orthographic case. Dourou and Courteille [4] start from the discretized functional over pairs (i, j) of adjacent pixels

$$E_{\mathrm{DC}} = \sum_{(i,j)\in\mathcal{N}} \left((z_j - z_i) - g_{ij}\right)^2 . \tag{14}$$

where g_{ij} is the expected gradient along the edge (i, j). It is derived from photometric normals. The discretized functional in Heep and Zell [5] (adapted to the orthographic case) is

$$E_{\mathrm{HZ}} = \sum_{(i,j)\in\mathcal{N}} \left((\boldsymbol{e}_z\cdot\boldsymbol{n}_{ij})\cdot(z_j - z_i) - (\boldsymbol{e}_{ij}\cdot\boldsymbol{n}_{ij})\right)^2 \tag{15}$$



(a) Gradient Formulation [4] **(b)** Normal Formulation [5] **(c)** Normal Formulation (Ours)

Fig. 1: Integration results for the pixel-based integration methods by [4] (a) and [5] (b) as well as ours (c). The gradient-based version exhibits the distortions known as the Gibbs phenomenon around the base of the sphere. These distortions are absent in the two normal-based approaches.

where $\boldsymbol{e}_z$ is the unit vector in the z direction, $\boldsymbol{e}_{ij} = (u_j - u_i, v_j - v_i, 0)^{\mathrm{t}}$ is the direction of the edge and $\boldsymbol{n}_{ij}$ is the average normal of pixels i and j. The latter, normal-based formulation can be seen as the pixel version of our mesh-based approach. Due to the relation

$$g_{ij} \approx \frac{\boldsymbol{e}_{ij} \cdot \boldsymbol{n}_{ij}}{\boldsymbol{e}_z \cdot \boldsymbol{n}_{ij}}\,, \tag{16}$$

the difference between the two approaches [4, 5] is essentially different edge weights. These edge weights put a big emphasis on regions where $n_z \approx 0$ in the gradient-based formulation. It is exactly these highly slanted regions, where the Gibbs phenomenon occurs, see the distortions around the sphere in Fig. 1. These distortions are completely absent in normal-based formulations.

4 Evaluation

Meshing First vs Pixel-based

In line with Fig. 2 from the main paper, we provide additional examples comparing the reconstruction accuracy where we roughly match the number of vertices of our meshing-first approaches to previous pixel-based integration methods. The female, male and dragon in Fig. 2 were used to create Fig. 6 in the main work. The human characters were purchased from 3D Scan Store, the Dragon is part of the Stanford 3D Scanning Repository [2].

In Fig. 3, we compare the integration error of pixel-based integration to our meshing-first approach when roughly matching the number of free variables, *i.e.* pixels or vertices respectively. The "Joint" is part of the dataset in [7]. The "SICILY. Syracuse. Second Democracy" coin by Frank McMains was licensed under CC BY 4.0. The "leaf1" was purchased from 3Dexport.

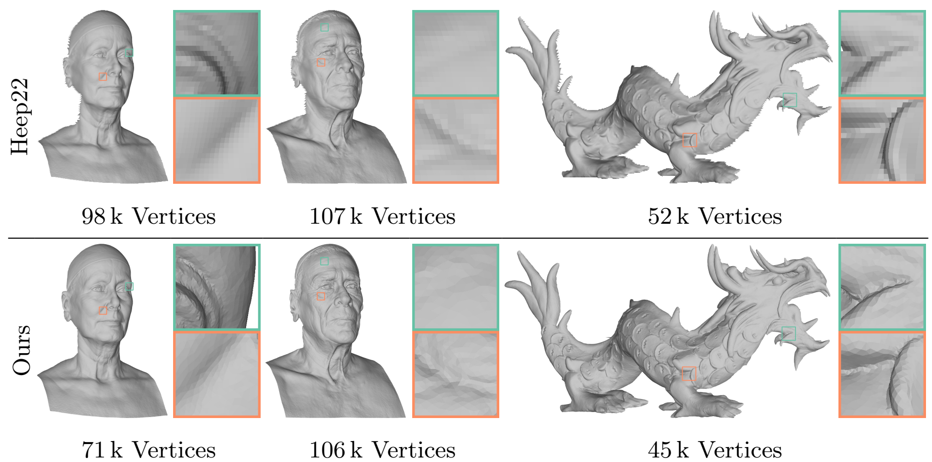


Fig. 2: Comparison between the results of our adaptive mesh integration and the pixel-based integration in [5]. We compare results with approximately the same number of vertices or pixels respectively. Flat shading was applied to visualize the structural differences between the regular pixel grid and our adaptive triangle mesh.

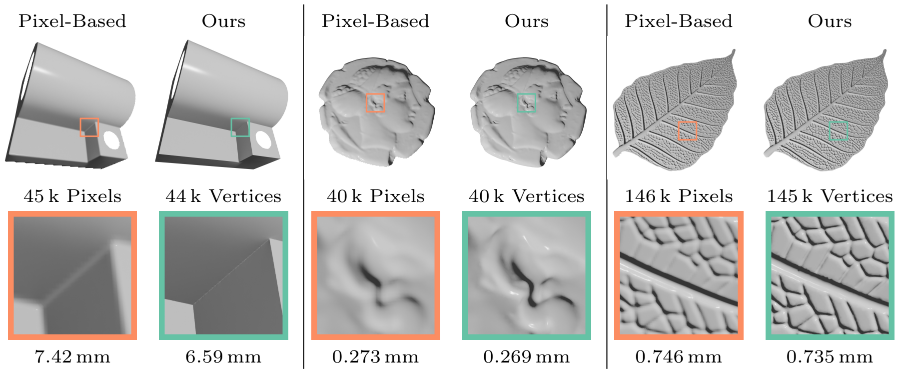


Fig. 3: Comparison of pixel-based integration [5] against our proposed meshing-first approach, where we roughly match the number of vertices to the number of pixels. Values in mm are the RMSEs of the test object. Our adaptive vertex placement can better capture the overall shape of the object and leads to better results both qualitatively and quantitatively.

4.1 Compression and Runtime

In Table 1, we reported the average compressions of our method for the low (1 mm), mid (0.3 mm) and high (0.1 mm) quality settings. We reported compressions for ground truth as well as computed normals [5,6]. In this section, we complement these numbers with visualizations of the created meshes, see Fig. 4 to Fig. 9.

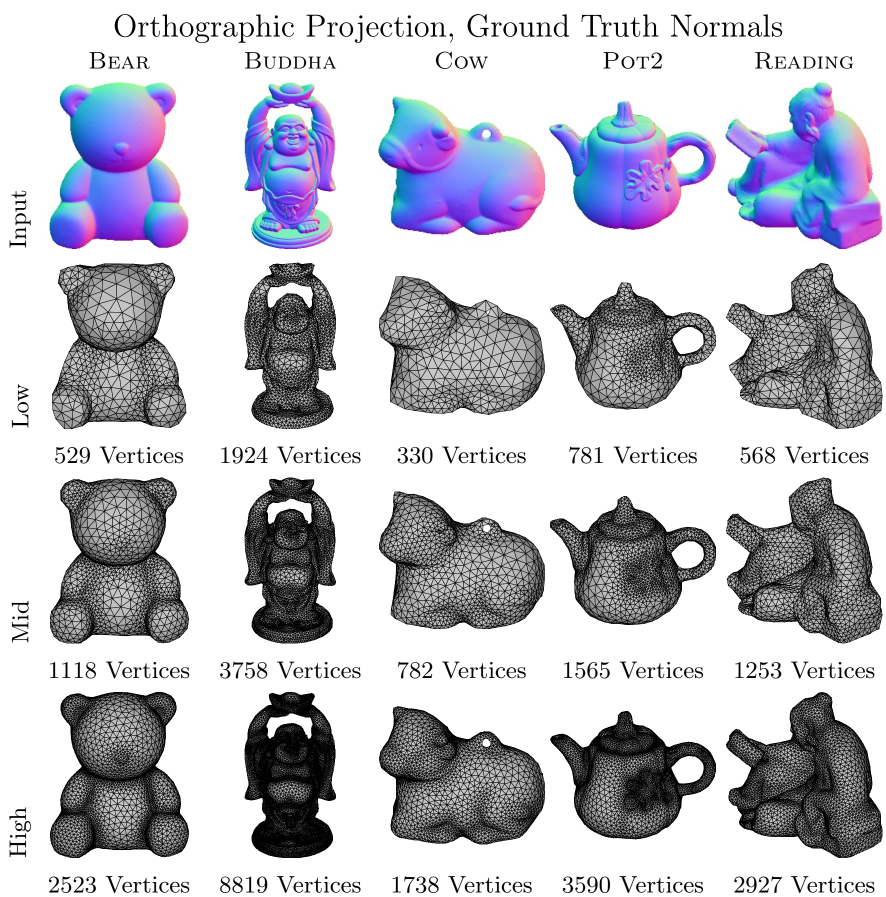


Fig. 4: Results for different quality settings of our adaptive screen-space meshing using ground truth normals and the orthographic projection. Wireframes are rendered as vector graphics for better examination.

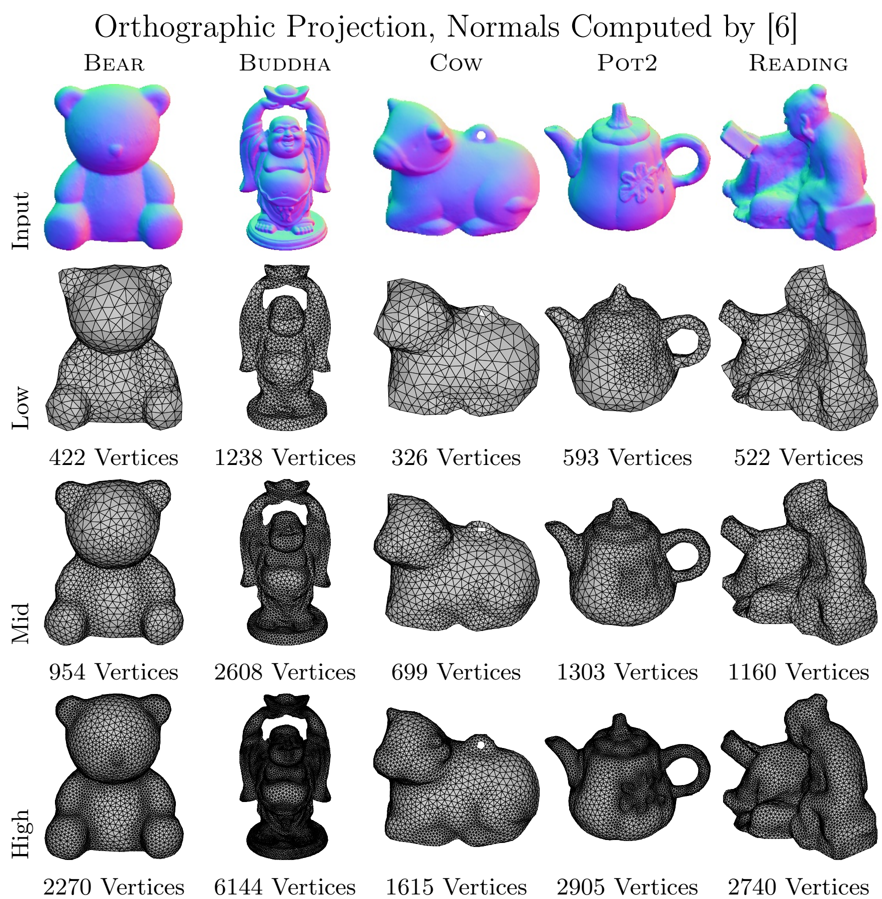


Fig. 5: Results for different quality settings of our adaptive screen-space meshing using normals calculated with [6] and the orthographic projection. Wireframes are rendered as vector graphics for better examination.

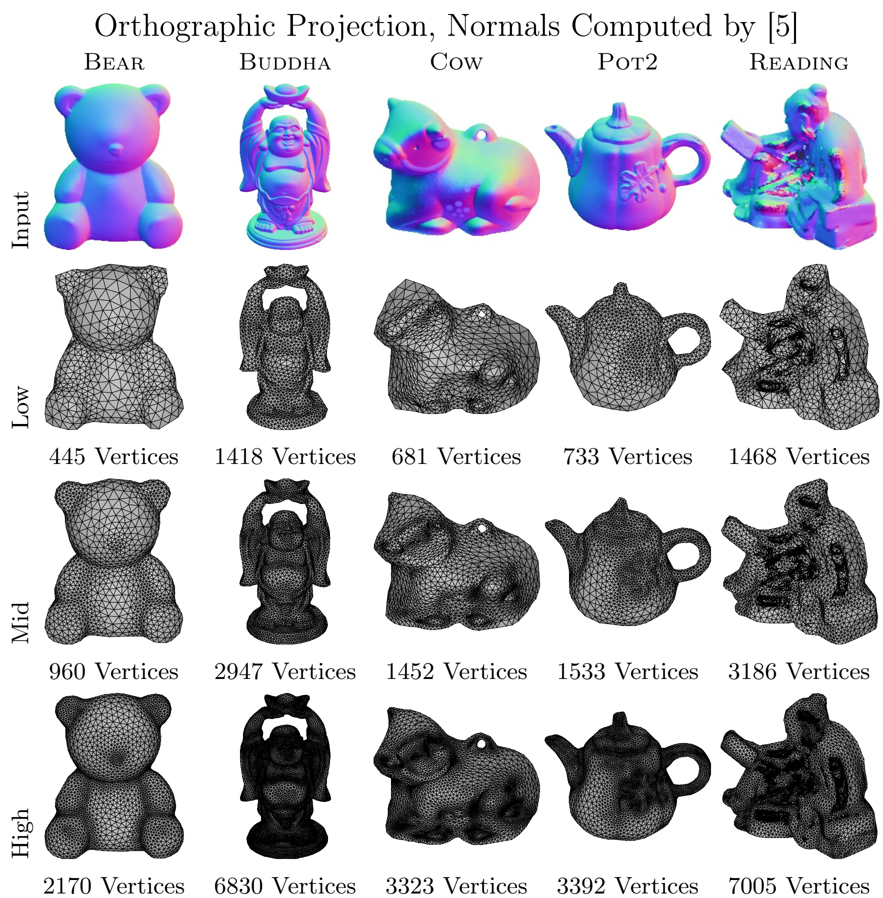


Fig. 6: Results for different quality settings of our adaptive screen-space meshing using normals calculated with [5] and the orthographic projection. The method [5] operates on the Lambertian surface assumption and clearly struggles with some of the more complex materials in DiliGenT-MV, *e.g.* the specular 'Reading' dataset. While incorrect normals affect meshing results, it does not break our meshing pipeline. Wireframes are rendered as vector graphics for better examination.

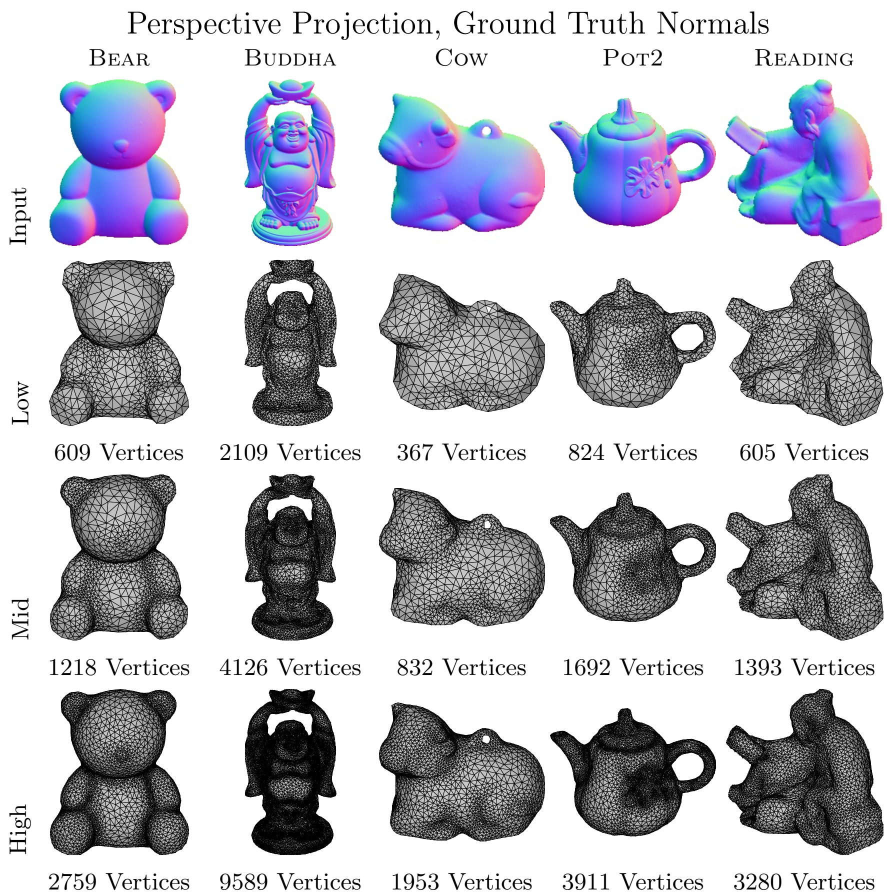


Fig. 7: Results for different quality settings of our adaptive screen-space meshing using ground truth normals and the perspective projection. Wireframes are rendered as vector graphics for better examination.

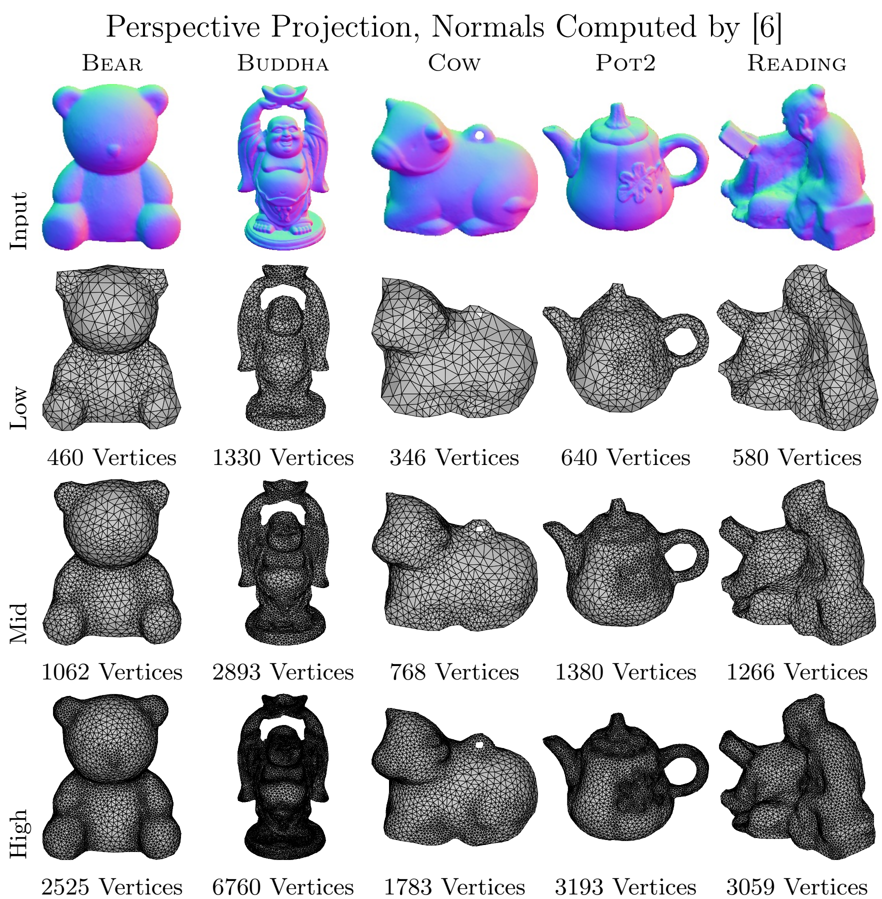


Fig. 8: Results for different quality settings of our adaptive screen-space meshing using normals calculated with [6] and the perspective projection. Wireframes are rendered as vector graphics for better examination.

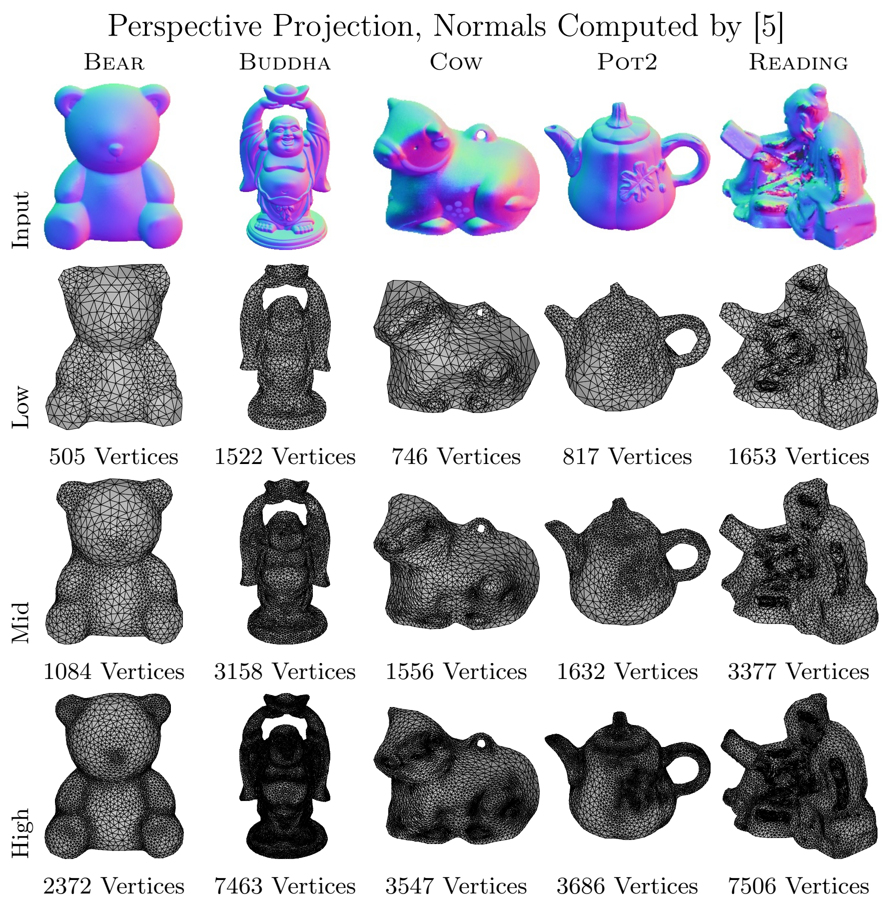


Fig. 9: Results for different quality settings of our adaptive screen-space meshing using normals calculated with [5] and the perspective projection. The method [5] operates on the Lambertian surface assumption and clearly struggles with some of the more complex materials in DiliGenT-MV, *e.g.* the specular 'Reading' dataset. While affecting meshing results, it does not break our meshing pipeline. Wireframes are rendered as vector graphics for better examination.

4.2 Benchmark Evaluation

In Table 2, we reported the RMSE after integration of our method compared to a range of previous works [1, 3, 5, 10]. In this section, we complement these numbers with absolute error maps (Figs. 10 and 11) and rotated views of the integrated surfaces (Figs. 12 and 13).

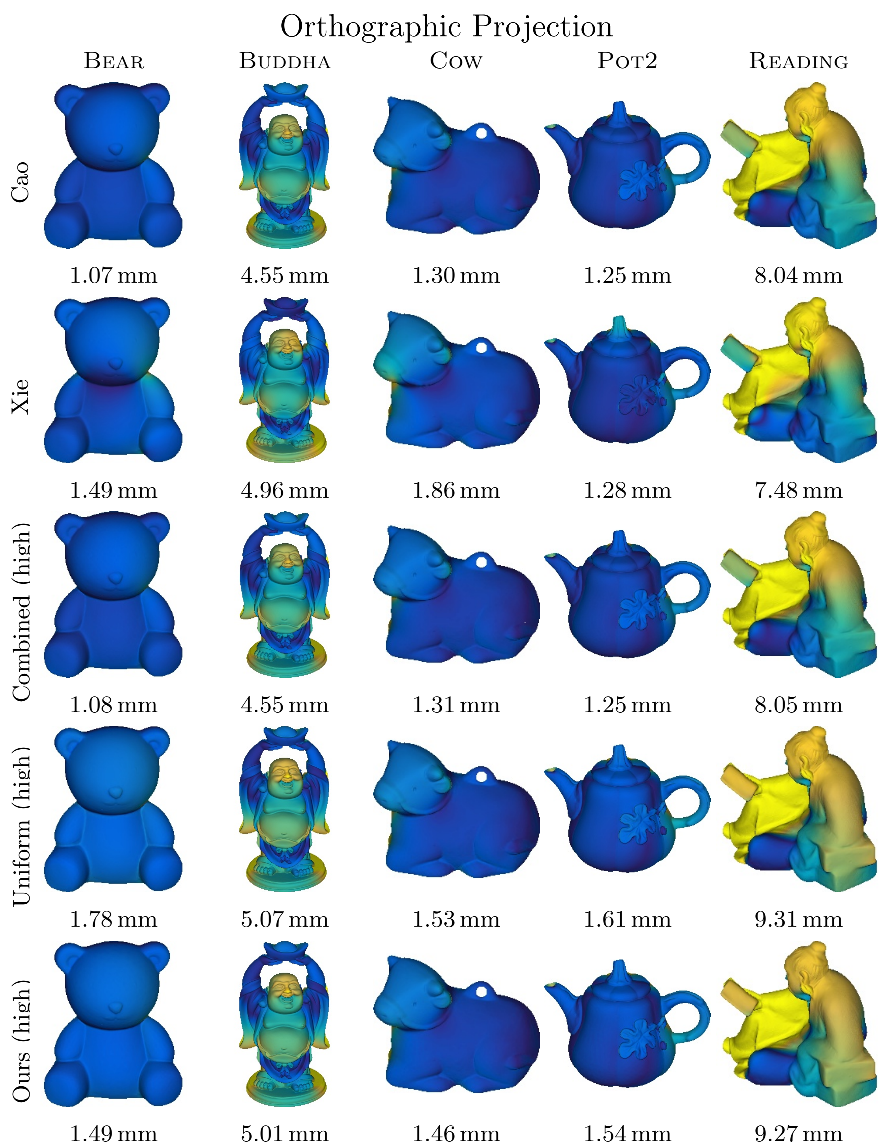


Fig. 10: RMSE and absolute error maps for the first view of the DiLiGenT-MV dataset in the orthographic projection.

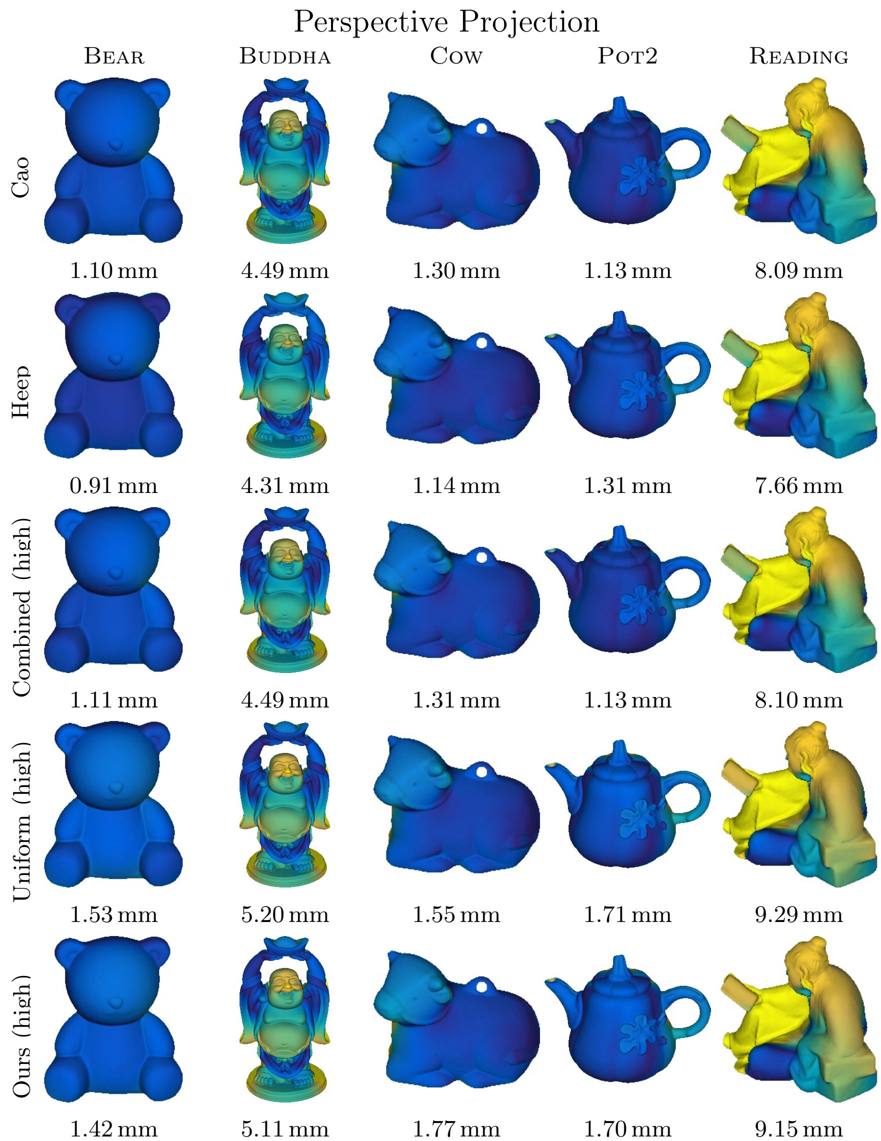


Fig. 11: RMSE and absolute error maps for the first view of the DiLiGenT-MV dataset in the perspective projection.

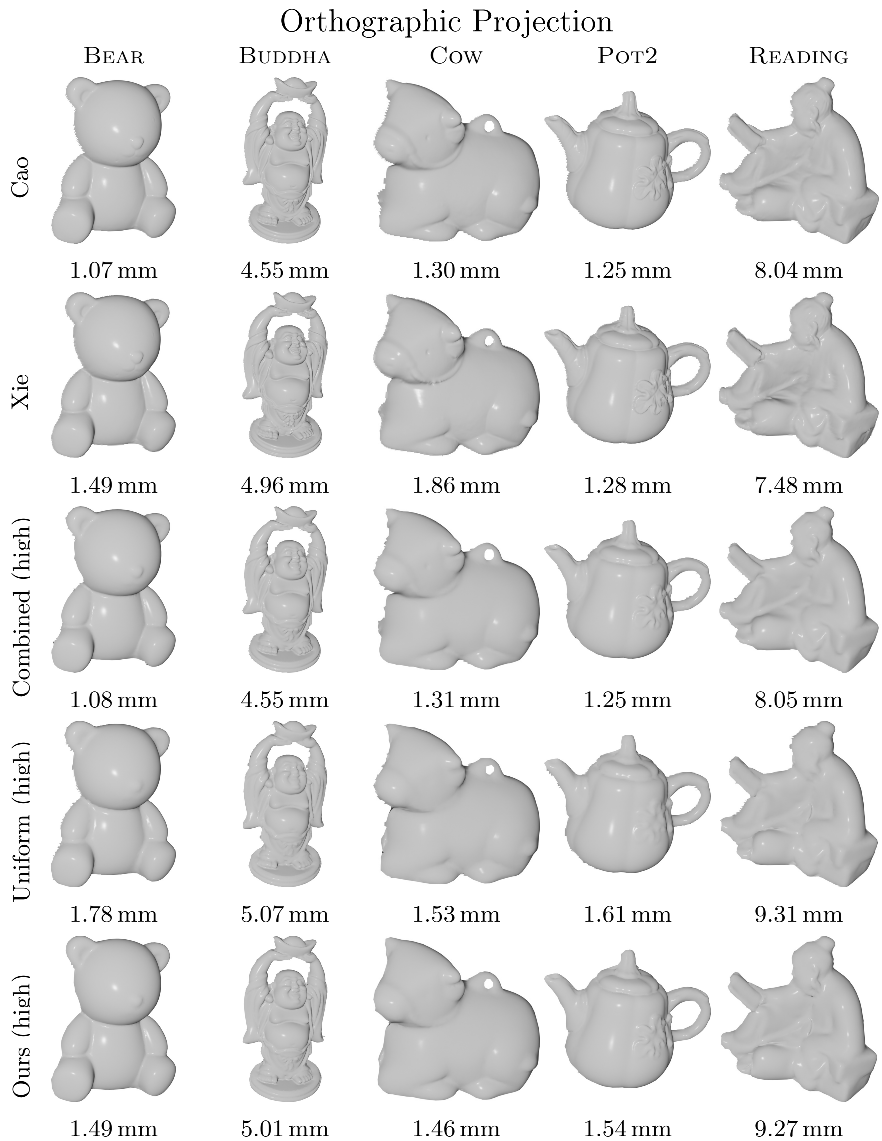


Fig. 12: Integrated surfaces for the first view of the DiLiGenT-MV dataset for the orthographic projection.

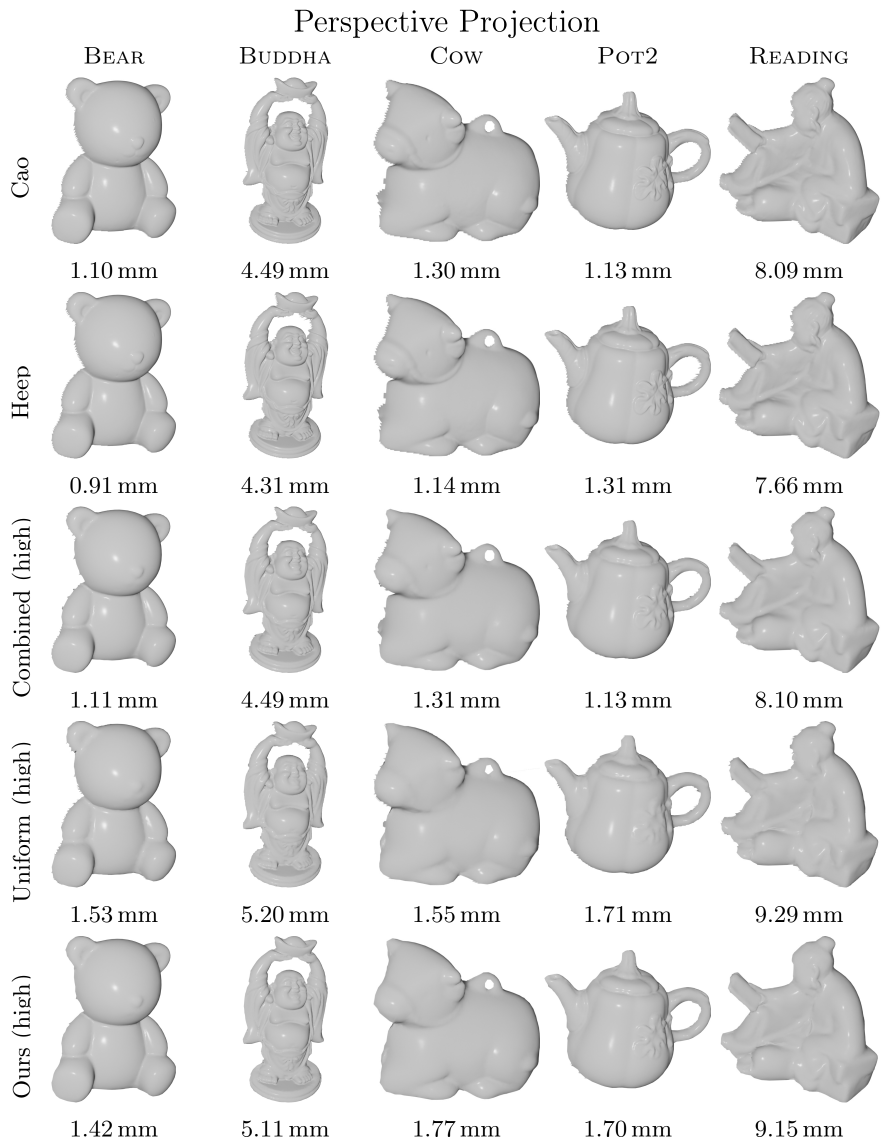


Fig. 13: Integrated surfaces for the first view of the DiLiGenT-MV dataset for the perspective projection.

4.3 Robustness towards Noise

To evaluate the impact of imperfect normals on the performance of our mesh-based integration, we added different levels of Gaussian noise to the ground truth normals maps, see Fig. 14. Our method can handle noise reliably. As expected, lower-quality input generates lower-quality output and diminishes surface details. High curvature regions with their small triangles are more affected while random perturbations tend to average out for bigger triangles in low curvature regions.

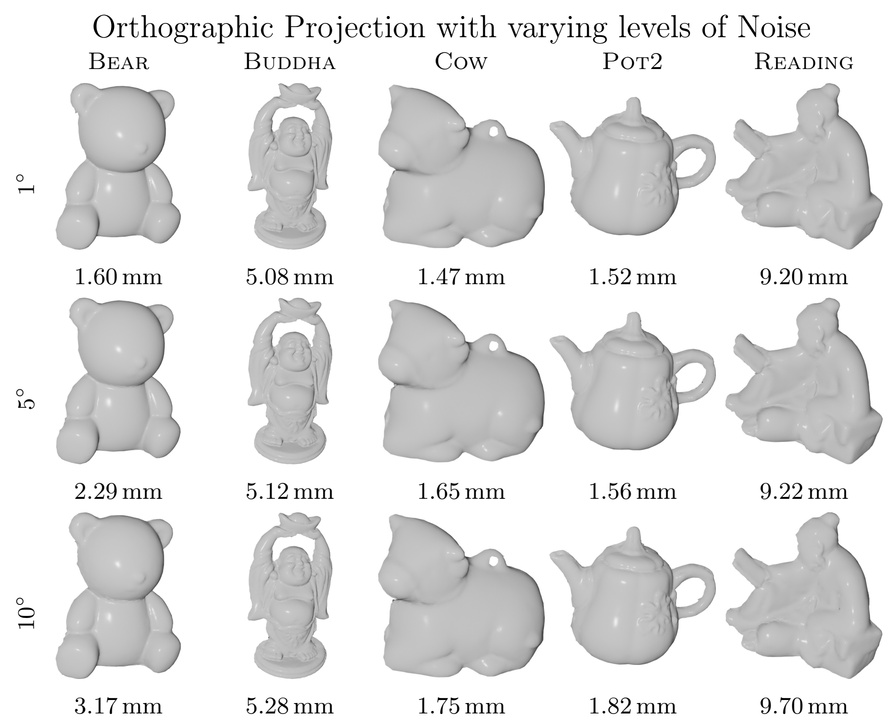


Fig. 14: Integrated surfaces for the first view of the DiLiGenT-MV dataset for three different levels of noise.

4.4 Ablation Study

For the investigation of the influence of the single user-parameter ϵ, we showed in the paper the remaining RMSE as a function of ϵ after fitting non-rigidly to ground truth. For completeness, we show the equivalent diagram for integrated meshes in Fig. 15. The linear connection between RMSE and ϵ still holds up after integration. However, additional errors that occur during integration lead to a bigger offset than the non-rigid fit to ground truth.

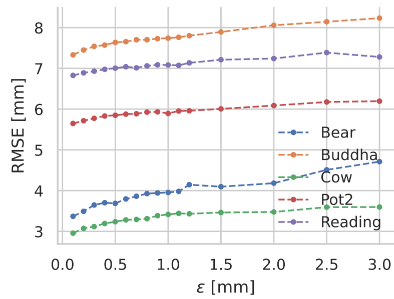


Fig. 15: Impact analysis of the user-parameter: RMSE after integration as a function of the user-parameter ϵ. The RMSE grows linearly with ϵ.

References

1. Cao, X., Shi, B., Fumio, O., Matsushita, Y.: Normal Integration via Inverse Plane Fitting With Minimum Point-to-Plane Distance. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2382–2391 (Jun 2021)
2. Curless, B., Levoy, M.: A volumetric method for building complex models from range images. In: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques. pp. 303–312 (1996)
3. Dunyach, M., Vanderhaeghe, D., Barthe, L., Botsch, M.: Adaptive remeshing for real-time mesh deformation. In: Eurographics 2013 - Short Papers. The Eurographics Association (2013)
4. Durou, J.D., Courteille, F.: Integration of a Normal Field without Boundary Condition. In: Proceedings of the First International Workshop on Photometric Analysis For Computer Vision-PACV 2007. pp. 8–p. INRIA (2007)
5. Heep, M., Zell, E.: ShadowPatch: Shadow Based Segmentation for Reliable Depth Discontinuities in Photometric Stereo. Computer Graphics Forum **41**(7), 635–646 (Oct 2022). `https://doi.org/10.1111/cgf.14707`
6. Li, M., Zhou, Z., Wu, Z., Shi, B., Diao, C., Tan, P.: Multi-view photometric stereo: A robust solution and benchmark dataset for spatially varying isotropic materials. IEEE Transactions on Image Processing **29**, 4159–4173 (2020)
7. Myles, A., Pietroni, N., Zorin, D.: Robust field-aligned global parametrization. ACM Trans. Graph. **33**(4), 135–1 (2014)
8. Pinkall, U., Polthier, K.: Computing discrete minimal surfaces and their conjugates. Experimental mathematics **2**(1), 15–36 (1993)
9. Vermolen, F.J., Segal, A.: On an integration rule for products of barycentric coordinates over simplexes in rn. Journal of Computational and Applied Mathematics **330**, 289–294 (2018)
10. Xie, W., Zhang, Y., Wang, C.C., Chung, R.C.K.: Surface-from-gradients: An approach based on discrete geometry processing. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2195–2202 (2014)