Seeing Through the Glass: Neural 3D Reconstruction of Object Inside a Transparent Container
Supplementary Material

Jinguang Tong[1,2], Sundaram Muthu[2], Fahira Afzal Maken[2], Chuong Nguyen[1,2], Hongdong Li[1]

[1]The Australian National University   [2]Data61, CSIRO

{jinguang.tong, chuong.nguyen, hongdong.li}@anu.edu.au

{sundaram.muthu, fahira.afzalmaken, chuong.nguyen}@data61.csiro.au

1. Details of dataset

We summarize our proposed dataset in Tab. 1. Some examples of our proposed dataset are shown in Fig. 5 and Fig. 6.

Dataset	N_{obj}	N_{view}	Resolution	Mask	Camera pose	GT model
w/box	13	60	800×800	box	GT	True
w/o box	13	60	800×800	object	GT	True
real	10	60	1728×1152	box	estimated	False

Table 1. Summary of the proposed dataset.

1.1. Synthetic dataset

Our synthetic dataset consists of 13 objects with various appearances and geometries, we list the link and license of each item in Tab. 3. The material of the transparent box is set to be a glass-like material with `Principled BSDF`[1] whose main parameters are shown in Tab. 2 and the background irradiance is set as $C^{out} = [0.8, 0.8, 0.8]$ across all the objects. While rendering, we render two images **w/ box** and **w/o box** separately by changing the visibility of the transparent box at each viewport. For the image mask used for training, we render the mask of the transparent box for w/ box dataset and the mask of the object for w/o box dataset. Overall, the two subsets share almost the same configurations except for the transparent box.

Parameter	Specular	Roughness	Sheen Tint	IOR	Transmission
Value	0.5	0.05	0.5	1.45	1.0

Table 2. Main parameters of the transparent material.

(a) Capture set-up.

(b) Edges.

(c) Intersections.

(d) Box geometry.

(e) Cropped image.

Figure 1. Process of real dataset calibration. (a) Data capture equipment. (b) Labeled edges of the box. (c) Intersections of the edges. (d) Calibrated geometry of the box. (e) Cropped image and box mask.

Item	License	URL
beetle	CC BY 4.0	https://sketchfab.com/3d-models/eastern-hercules-beetle-a6f2a6fe7a3c4d19831f1bd79d664dcc
box	CC BY 4.0	https://sketchfab.com/3d-models/tollkiste-00fa12d17fe04dc39e9275164f5339c1
butterfly	CC BY 4.0	https://sketchfab.com/3d-models/common-lime-butterfly-smoothie-3d-078da6e1f86c4d2ca874e3f4381becb8
coral	CC BY-NC-SA 4.0	https://sketchfab.com/3d-models/pocillopora-eydouxi-b2a7aef9f4814de09d42546c1a974cd0
coral2	CC BY 4.0	https://sketchfab.com/3d-models/pocillopora-meandrina-3bacf518167b46d78fef48ef9e0b7019
dinosaur	CC BY 4.0	https://sketchfab.com/3d-models/squelette-entier-de-felin-espece-inconnue-9b7ae513a2294711a9e083de1a4dcae5
goku	CC BY 4.0	https://sketchfab.com/3d-models/goku-3d-scan-dragon-ball-3db326f4b41f4a00bbd116fad7f4d09e
insect	CC BY 4.0	https://sketchfab.com/3d-models/cicindela-campestris-b13a625c2e3b4b6aa26a27711a0cac39
insect2	CC BY 4.0	https://sketchfab.com/3d-models/prosopocoilus-savagei-b04ad901d4344ccf8ec6893a6b36c27c
lobster	CC0 1.0	https://sketchfab.com/3d-models/panulirus-longipes-d7cc2b585ad943ef81274e014e6f745b
shiba	CC BY 4.0	https://sketchfab.com/3d-models/free-toon-shiba-inu-3d-model-ae0a62c3d588496fa9d5e516fcb6786c
statuette	TurboSquid standard	https://www.turbosquid.com/3d-models/3d-statuette-sheep-barrel-model-1335035
vase	CC BY 4.0	https://sketchfab.com/3d-models/vase-951156fb0db74176af08501bdb4b5acd

Table 3. Dataset source and license information.

1.2. Real dataset

The real dataset is captured with 10 insect specimens. As the set-up shown in Fig. 1a, the insect specimen is roughly placed at the center of a printed Charuco board [1]. We take photos of the specimen at different pitch angles while the board rotates uniformly on a display stand. Camera intrinsics and extrinsics are then calibrated with the Charuco board. As for the pose of the transparent box, we calibrate it manually, as shown in Fig. 1b-e. Firstly, we pick out the image with the lowest re-projection error during calibration and manually label the edges of the transparent box. Then, we calculate the 2D coordinates of the intersections of these lines. With known camera pose, the 2D coordinates can be lifted up to 3D by ray-plane intersecting. Finally, under the assumption of a cuboid-shaped box, we get the 3D geometry of the box as well as the relative pose between the box and the calibration board. The origin of the world coordinate system is transferred to the center of the box with the axis aligned and we further scale the poses so that the box lies in a unit sphere. Fig. 1d shows that our calibration result aligns with the box well. After calibration, we back-project the geometry of the box to each image to get the mask of the transparent box and center-crop the image to remove redundant boundaries.

2. Additional Experimental Details

2.1. Additional Implementation Details

Scene parameters. Our scene is assumed to be in the air. In order to do ray tracing, we set the Index of Refraction (IoR) of different mediums as $\lambda_{air} = 1.00029$, $\lambda_{syn} = 1.45$ and $\lambda_{real} = 1.5$, respectively.

Post process As mentioned in Sec. 4.3 in the main paper, we post-process the output mesh from w/ box dataset automatically. To be specific, we divide the mesh into several connected clusters and only retrain the biggest cluster with the most vertices. An example of post-processing is shown in Fig. 2. Since we are comparing with the results from w/o box dataset, it's a reasonable process that can exclude the influence of boundary effects and allow us to evaluate the quality of object reconstruction better.

(a) Before process. (b) After process.

Figure 2. Example of mesh post-processing. We only remove the trivial facets without any further modificaiton.

2.2. Baselines

COLMAP [4] We use the officially provided command line interface (CLI) of COLMAP. For synthetic dataset, we run COLMAP with ground truth pose and for real dataset, we use the calibrated camera pose. We run following CLI commands [2] to get a dense point cloud: $(1)feature\_extractor$, $(2)exhaustive\_matcher$, $(3)point\_triangulator$, $(4)patch\_match\_stereo$ and $(5)stereo\_fusion$. With the dense point cloud, we use the screened Poisson Surface Reconstruction (sPSR) [2] to generate the corresponding mesh.

IDR [6] We run IDR with the official code release [3]. We train IDR with our w/ box dataset using the same configuration as the original paper.

NeuS We use the official code release [4] of NeuS. For a fair comparison with our method, we also replace the activation function in NeuS with SIREN [5] and train NeuS with a larger batch size of 1024 for 200k iterations. We train NeuS with an image mask as the better performance shown in the paper.

2.3. Additional ablation study

In this section, we show some extended experiments of the ablation study in the main paper.

Influence of transmittance loss. We further analyze the influence of transmittance loss as an extend to the main paper. In Fig. 3, we visualize the evaluation metrics Chamfer-L_1 and the number of failure cases w.r.t. the weight of transmittance loss. From the results, we know that the transmittance loss can boost the performance of ReNeuS by reducing the failure cases from 2 to 0. Further, we adopt the weight of transmittance loss $\lambda_1 = 0.1$ as a good measure of the dataset sparsity.
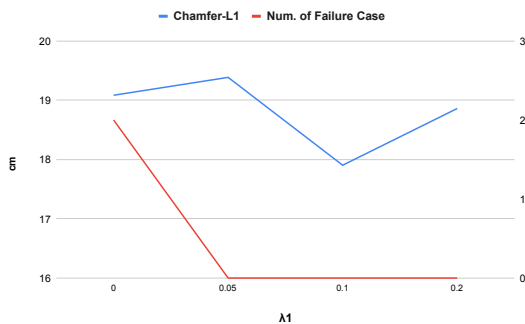
Figure 3. We show the influence of transmittance loss. Results show that the loss can promote the effectiveness of ReNeuS and we find $\lambda_1 = 0.1$ to be a balance between sparsity and reconstruction quality on the proposed dataset.

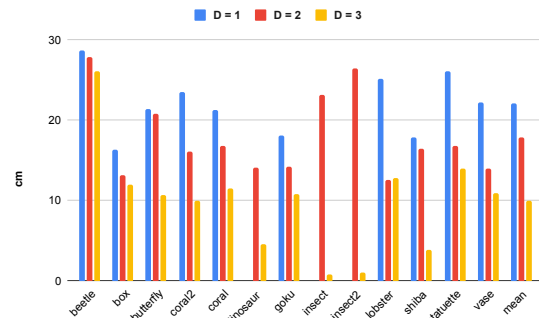**Influence of ray tracing depth.** We evaluate the per-

Figure 4. We evaluate the performance of ReNeuS at different tracing depth. Our method generally performs better with deeper ray tracing depth showing that the proposed hybrid rendering strategy models the scene well.

formance of ReNeuS w.r.t. the recursive depth D_{re} of ray tracing in Fig. 4. We alternate the depth from 1~3, among which we adopt the naive solution $NeuS^+$ in the main paper as $D_{re} = 1$. Note that for $D_{re} = 3$, we have to half the training batch size due to memory limitation and double total iterations correspondingly. We don't trace "deeper" because the contribution of those rays can be quite limited due to light attenuation [**?**] and absorption [3]. The results show the reconstruction quality is consistently promoted with deeper ray tracing which indicates our hybrid rendering strategy fit the complex scene well.

2.4. Additional qualitative comparison

In this section, we show additional qualitative results on both datasets. Figure 7 shows the comparison with baseline methods on synthetic dataset. Even working with w/ box dataset and no object mask, our method produces comparable results with IDR which works on w/o box dataset and with an object mask. In Fig. 8, we demonstrate reconstruction results on real dataset. We find that our ReNeuS is the only one which works properly on such challenging scenes.

References

[1] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6):2280–2292, 2014. 2

[2] Michael Kazhdan and Hugues Hoppe. Screened poisson surface reconstruction. *ACM TOG*, 32(3):1–13, 2013. 3

[3] N. Max. Optical models for direct volume rendering. *IEEE TVCG*, 1(2):99–108, 1995. 3

[4] Johannes L. Schonberger and Jan-Michael Frahm. Structure-From-Motion Revisited. In *CVPR*, pages 4104–4113, 2016. 3

[5] Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit Neural Representa-

tions with Periodic Activation Functions. In *NeurIPS*, volume 33, pages 7462–7473, 2020. 3

[6] Lior Yariv, Yoni Kasten, Dror Moran, Meirav Galun, Matan Atzmon, Basri Ronen, and Yaron Lipman. Multiview neural surface reconstruction by disentangling geometry and appearance. In *NeurIPS*, volume 33, pages 2492–2502, 2020. 3

	w/box	w/o box	w/box	w/o box	w/box	w/o box
beetle						
box						
butterfly						
coral						
coral2						
dinosaur						
goku						

Figure 5. Examples of synthetic dataset.
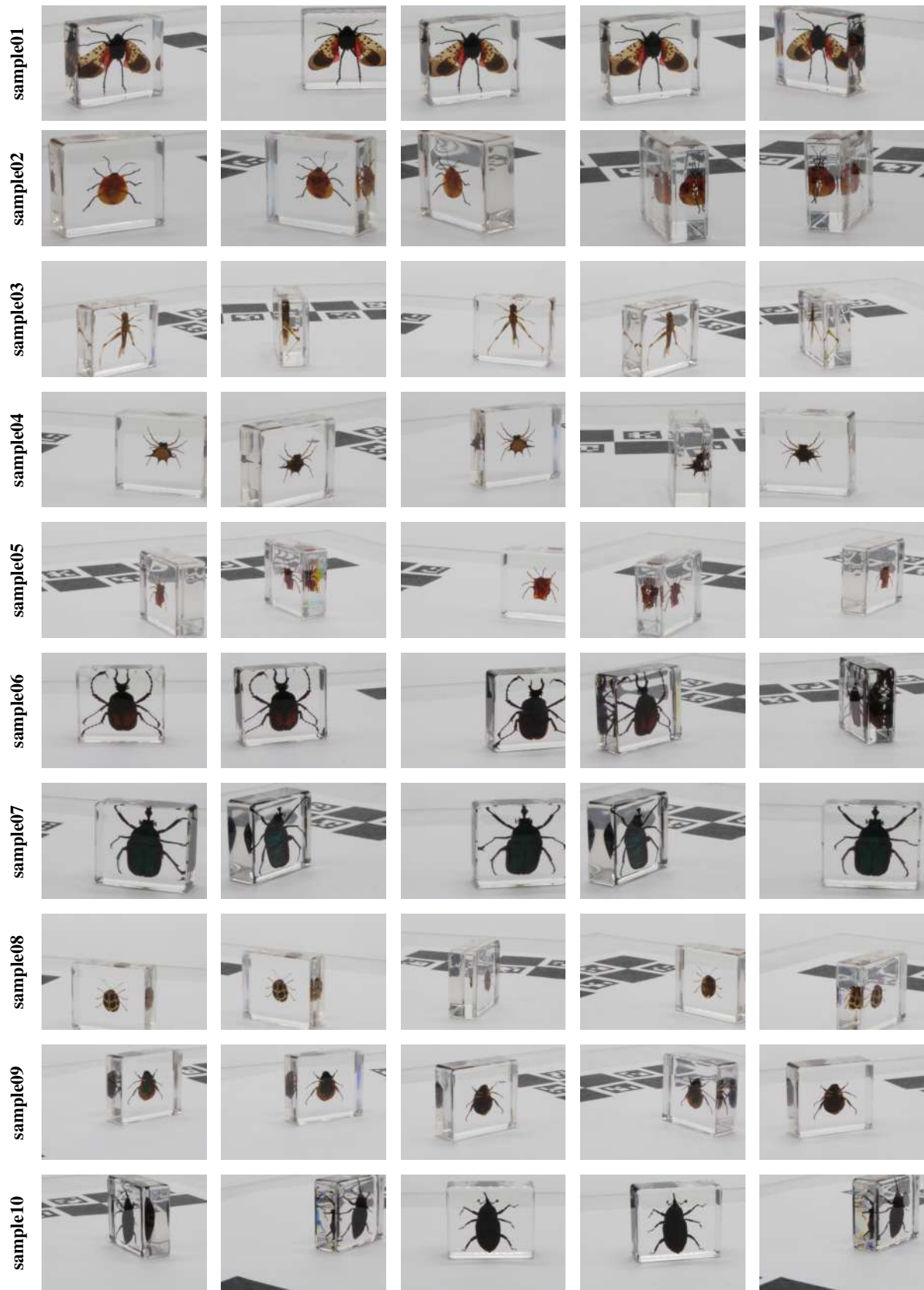
Figure 6. Examples of real dataset

	w/o box			w/box		
GT view	COLMAP	IDR	NeuS	COLMAP	NeuS	**Ours**

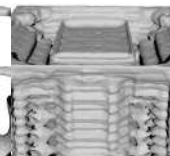**w/o box**   **w/box**
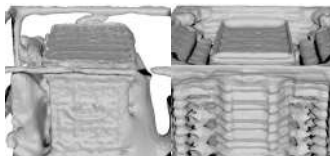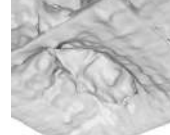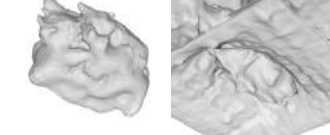
insect2   lobster   shiba   statuette   vase
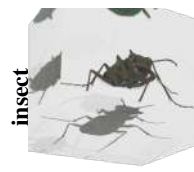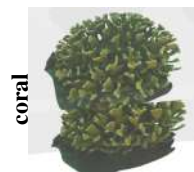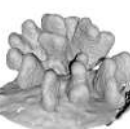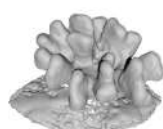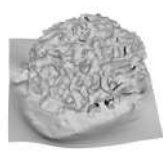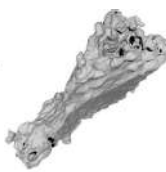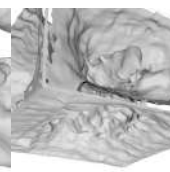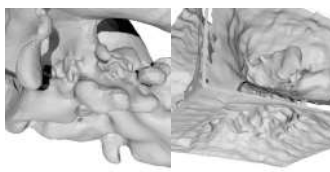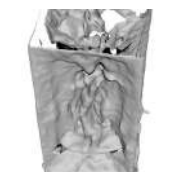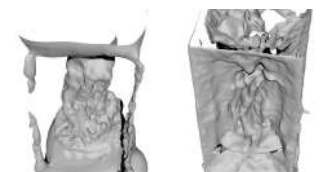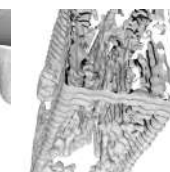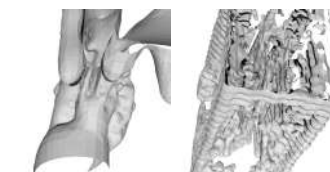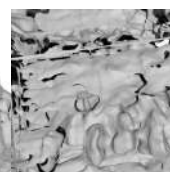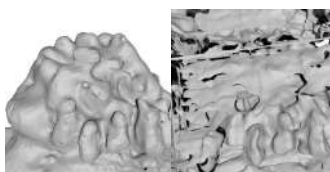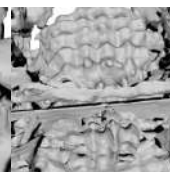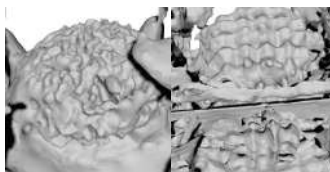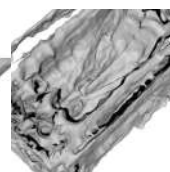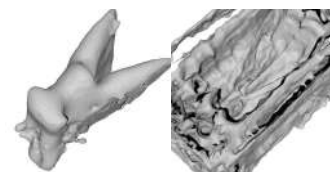
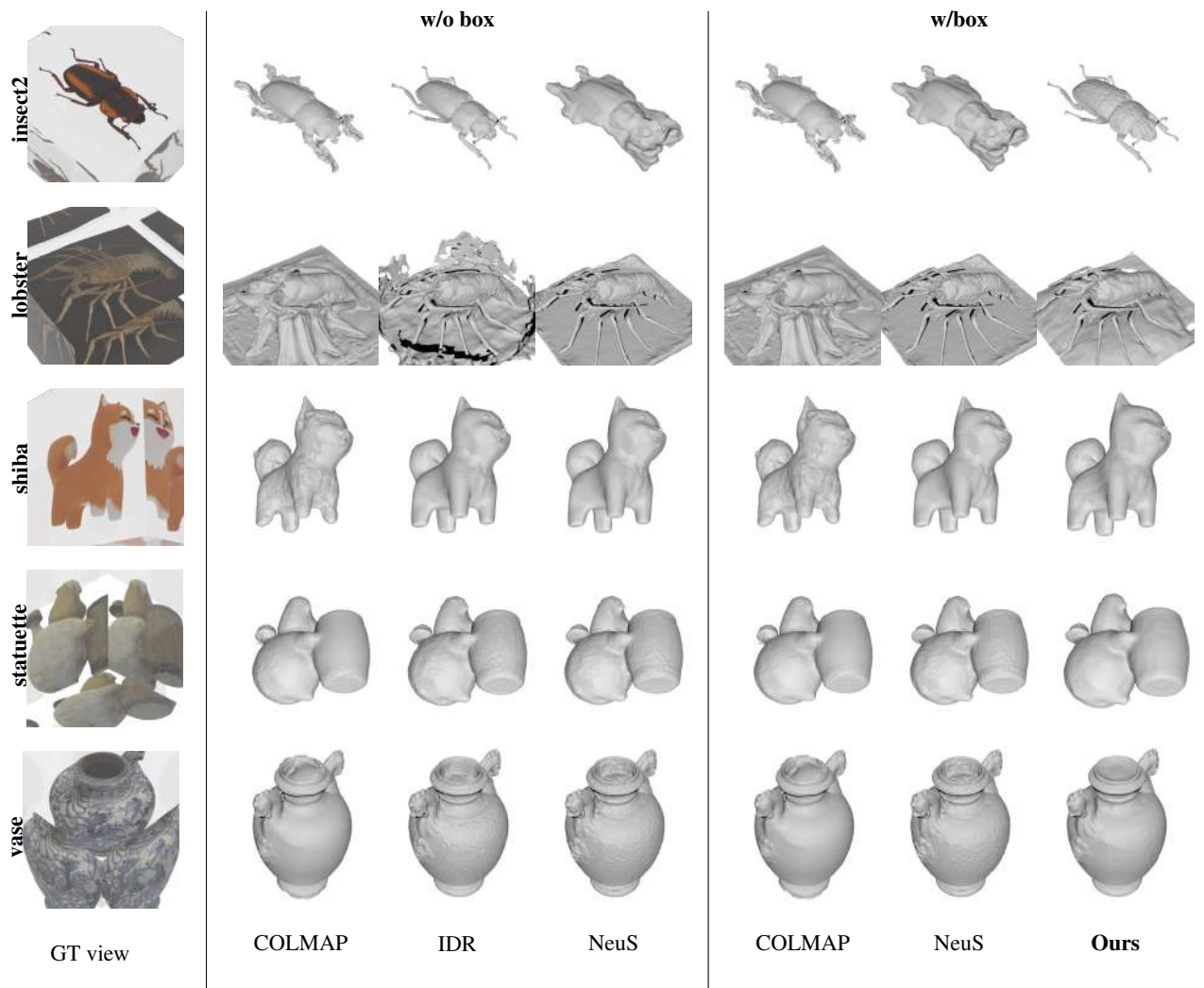GT view   COLMAP   IDR   NeuS   COLMAP   NeuS   **Ours**

Figure 7. Qualitative comparison on synthetic dataset.

Figure 8. Qualitative comparison on real dataset.