Supplementary Document for *Active View Selection with Perturbed Gaussian Ensemble for Tomographic Reconstruction*

This supplementary document is organized as follows. We first detail the datasets and implementation specifics in Secs. A and B, respectively. Subsequent sections present extended experimental evaluations, including additional quantitative comparisons in Sec. C and comprehensive qualitative visual results in Sec. D. Finally, Sec. E concludes with a discussion on the broader positive and negative impacts of our proposed approach. *Our source code is also attached.*

A More Details of Datasets

A.1 Synthetic Dataset

The synthetic dataset used in our work consists of 15 cases across three categories: human organs (chest, foot, head, jaw, and pancreas), plants and animals (beetle, bonsai, broccoli, kingsnake, and pepper), and artificial objects (backpack, engine, mount, present, and teapot). The chest scan is from LIDC-IDRI [1]; the pancreas scan is from Pancreas-CT [12]; the broccoli and pepper scans are from X-Plant [14]; and the rest are from SciVis [8]. Following [16], the raw data densities are normalized to $[0, 1]$, and the ground truth (GT) volumes are resized to a resolution of $256 \times 256 \times 256$. We use the GT projections at a resolution of 512×512, rendered with DiffDRR [4] from arbitrary poses. All GT volumes are shown in Fig. 1.

A.2 Real-World Dataset

The real-world dataset [13] consists of three objects: a pine cone [9], a seashell [6], and a walnut [10]. Each case contains 721 projections captured along an object-centric orbit. Because ground truth volumes are unavailable, we employ the FDK algorithm [3] using all available views to generate pseudo-ground-truth (pseudo-GT) volumes. From these, we render pseudo-GT projections at a resolution of 560×560 from arbitrary poses using DiffDRR [4]. The target volumes are reconstructed at a resolution of $256 \times 256 \times 256$.

B More Implementation Details

Our pipeline is implemented in PyTorch [11] 1.12.1 and CUDA 11.6. Following [16], the initial learning rates for position, density, scale, and rotation in our experiments are set to 0.0002, 0.01, 0.005, and 0.001, respectively, decaying exponentially to 0.1 times their initial values by the end of training. The loss weights are set to $\lambda_1 = 0.25$ and $\lambda_2 = 0.05$. We initialize 50k Gaussians with a

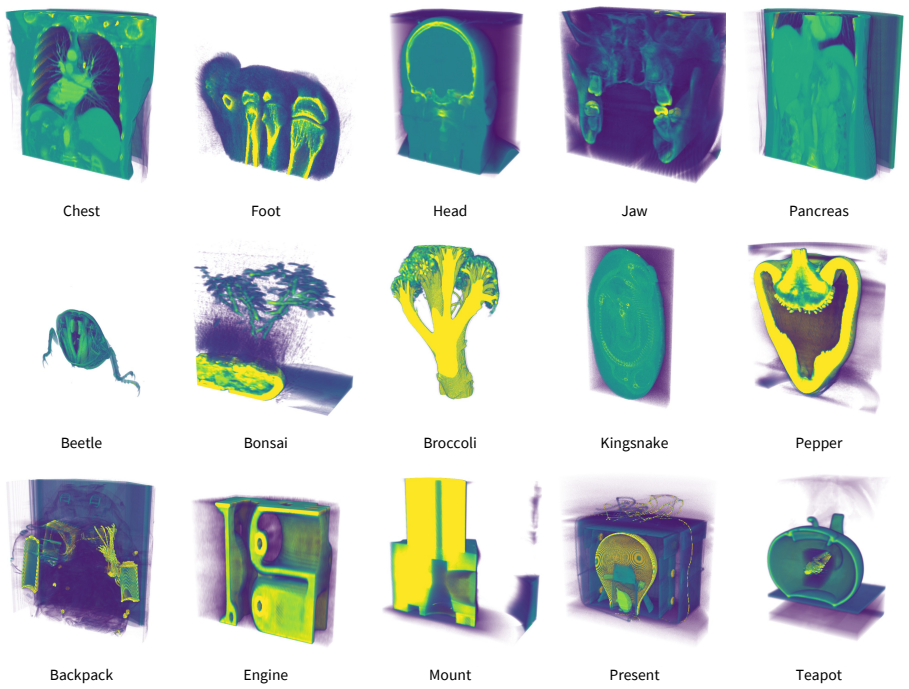

Fig. 1: Visualization of ground truth volumes from the synthetic dataset. A longitudinal cross-section is shown for each object.

density threshold of $\tau = 0.05$ and a scaling term of $k = 0.15$. The Total Variation (TV) volume size is set to $D = 32$. Adaptive density control is applied from 500 to 15,000 iterations, using a gradient threshold of 0.00005.

C More Quantitative Results

We also evaluate our method's performance on novel view synthesis using the real-world dataset. The quantitative results are reported Tab. 1. Our approach achieves the best rendering quality under both protocols.

D More Visualization Results

While the main text provides several representative cases for the visual comparison of 3D volume reconstructions, here we present comprehensive qualitative results for all evaluated cases across both synthetic and real-world datasets. Specifically, Figs. 2 to 4 illustrate the reconstruction results for human organs, plants and animals, and artificial objects from the synthetic dataset, respectively. Furthermore, Fig. 5 details the visual comparisons on the real-world dataset, benchmarking our method against all baseline approaches.

Method	24-view		36-view	
	PSNR↑	**SSIM↑**	**PSNR↑**	**SSIM↑**
Random	42.992	0.985	44.965	0.989
FPS	43.073	0.985	45.171	0.989
TOPIQ [2]	43.121	0.986	45.644	0.990
MANIQA [15]	<u>43.476</u>	<u>0.986</u>	45.891	<u>0.990</u>
MUSIQ [7]	43.293	0.985	45.494	0.990
FisherRF [5]	43.179	0.985	<u>46.022</u>	0.990
Ours	**43.546**	**0.987**	**46.046**	**0.991**

Table 1: Quantitative comparison of novel view synthesis on the real-world dataset using two different protocols. The **best** and <u>second-best</u> results are highlighted.

E Potential Impacts

E.1 Potential Positive Impacts

The primary positive societal impact of our work lies in the advancement of healthcare and occupational safety. By enabling high-fidelity tomographic reconstructions from a highly constrained number of X-ray projections, our active view selection framework significantly reduces the ionizing radiation exposure required for CT scans. This directly mitigates the associated health risks for patients, particularly in clinical scenarios requiring repetitive or continuous imaging, such as intraoperative CT and image-guided radiotherapy. Furthermore, by autonomously determining the most informative scanning trajectories, our method reduces the dependency on predefined, rigid scanning orbits, thereby enhancing the efficiency and flexibility of robotic non-destructive industrial inspection and scientific research.

E.2 Potential Negative Impacts

Conversely, as with many AI-driven medical imaging technologies, our framework carries potential negative impacts and clinical risks that must be carefully managed. A primary concern is the potential for over-reliance by medical practitioners on the progressively reconstructed volumetric data. In extreme sparse-view settings, if the initial random views completely miss a highly localized but critical anatomical anomaly, the model may fail to exhibit high structural variance in that specific area. Consequently, the active selection agent might not allocate subsequent views to explore that hidden region, leading to a confident but anatomically incomplete reconstruction.

Code Usage

For detailed instructions, please refer to `./code/README.md`

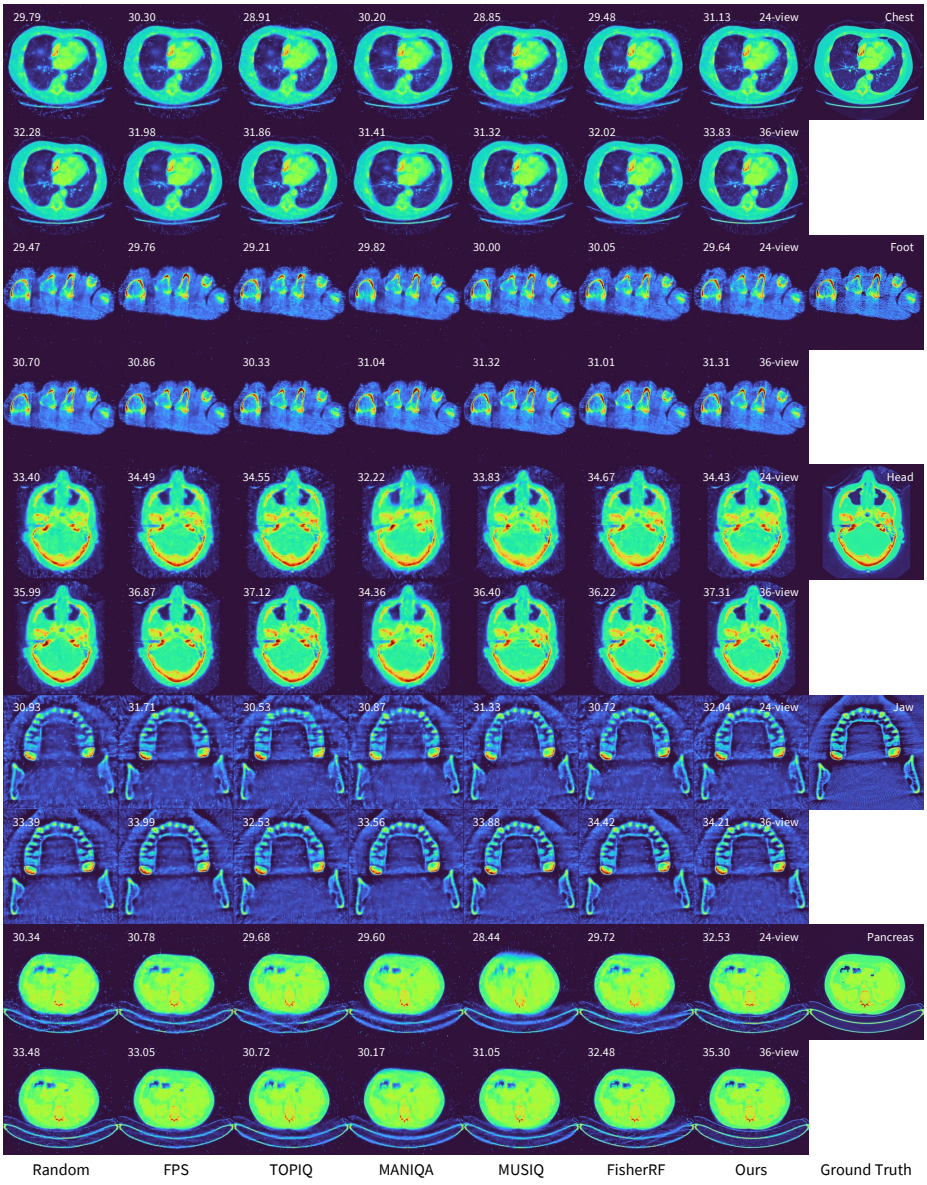

Fig. 2: Visual comparison of 3D volumes reconstructed under various view selection strategies across five human organ cases (chest, foot, head, jaw, and pancreas). The 3D PSNR↑ (dB) value for each case is displayed at the top-left corner of the image.

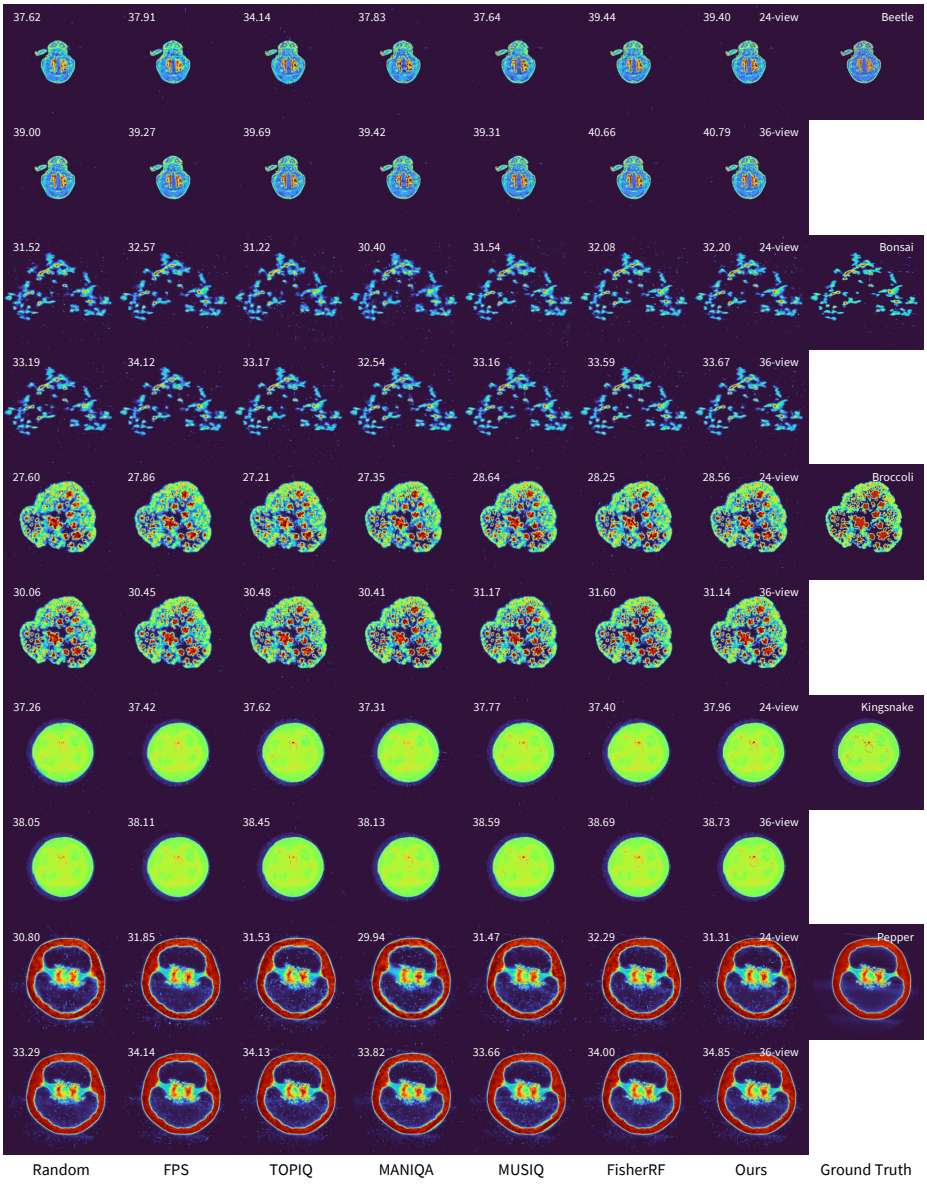

Fig. 3: Visual comparison of 3D volumes reconstructed under various view selection strategies across five plant/animal cases (beetle, bonsai, broccoli, kingsnake, and pepper). The 3D PSNR↑ (dB) for each case is displayed at the top-left corner of the image.

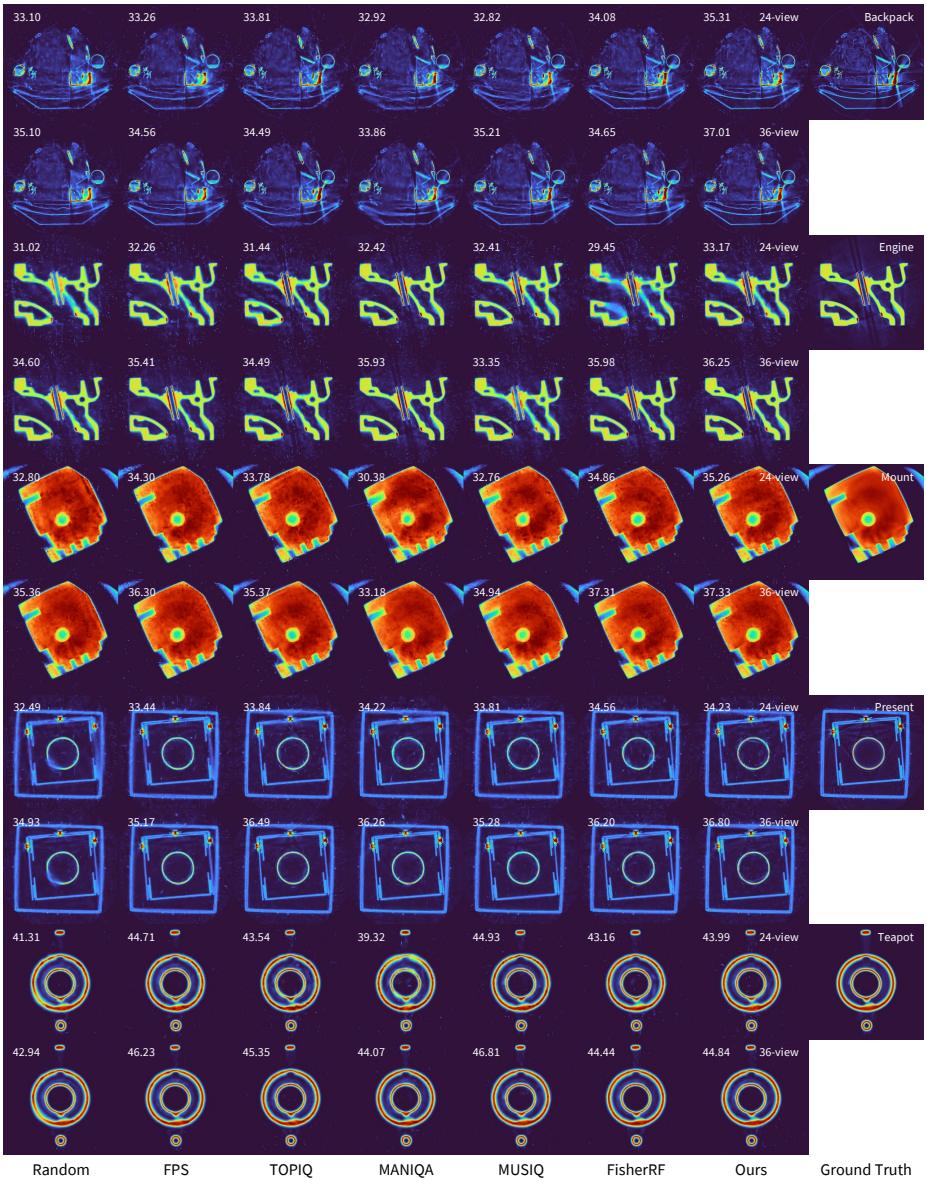

Fig. 4: Visual comparison of 3D volumes reconstructed under various view selection strategies across artificial object cases (backpack, engine, mount, present, and teapot). The 3D PSNR↑ (dB) for each case is displayed at the top-left corner of the image.

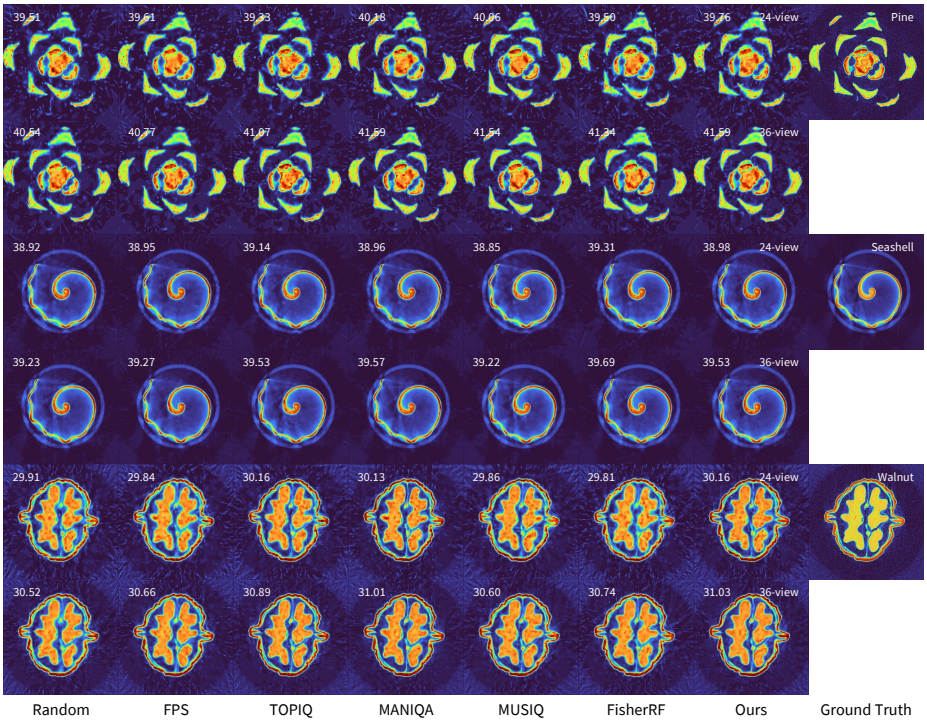

Fig. 5: Visual comparison of 3D volumes reconstructed under various view selection strategies across three real-world cases (pine, seashell, and walnut). The 3D PSNR↑ (dB) value for each case is displayed at the top-left corner of the image.

References

1. Armato III, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A., et al.: The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. Medical physics **38**(2), 915–931 (2011)
2. Chen, C., Mo, J., Hou, J., Wu, H., Liao, L., Sun, W., Yan, Q., Lin, W.: Topiq: A top-down approach from semantics to distortions for image quality assessment. TIP **33**, 2404–2418 (2024)
3. Feldkamp, L.A., Davis, L.C., Kress, J.W.: Practical cone-beam algorithm. Journal of the Optical Society of America A **1**(6), 612–619 (1984)
4. Gopalakrishnan, V., Golland, P.: Fast auto-differentiable digitally reconstructed radiographs for solving inverse problems in intraoperative imaging. In: Workshop on Clinical Image-Based Procedures. pp. 1–11. Springer (2022)
5. Jiang, W., Lei, B., Daniilidis, K.: Fisherrf: Active view selection and mapping with radiance fields using fisher information. In: ECCV. pp. 422–440. Springer (2024)
6. Kamutta, E., Mäkinen, S., Meaney, A.: Cone-Beam Computed Tomography Dataset of a Seashell (Aug 2022). https://doi.org/10.5281/zenodo.6983008
7. Ke, J., Wang, Q., Wang, Y., Milanfar, P., Yang, F.: Musiq: Multi-scale image quality transformer. In: ICCV. pp. 5148–5157 (2021)
8. Klacansky, P.: Open scivis datasets (December 2017), https://klacansky.com/open-scivis-datasets/
9. Meaney, A.: Cone-Beam Computed Tomography Dataset of a Pine Cone (Aug 2022). https://doi.org/10.5281/zenodo.6985407
10. Meaney, A.: Cone-beam computed tomography dataset of a walnut (Aug 2022). https://doi.org/10.5281/zenodo.6986012
11. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al.: Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems **32** (2019)
12. Roth, H., Farag, A., Turkbey, E.B., Lu, L., Liu, J., Summers, R.M.: Data from pancreas-ct (2016)
13. Society, T.F.I.P.: X-ray tomographic datasets (2024), https://fips.fi/category/open-datasets/x-ray-tomographic-datasets/
14. Verboven, P., Dequeker, B., He, J., Pieters, M., Pols, L., Tempelaere, A., Van Doorselaer, L., Van Cauteren, H., Verma, U., Xiao, H., et al.: www. x-plant. org-the ct database of plant organs. In: 6th Symposium on X-ray Computed Tomography: Inauguration of the KU Leuven XCT Core Facility, Location: Leuven, Belgium (2022)
15. Yang, S., Wu, T., Shi, S., Lao, S., Gong, Y., Cao, M., Wang, J., Yang, Y.: Maniqa: Multi-dimension attention network for no-reference image quality assessment. In: CVPR. pp. 1191–1200 (2022)
16. Zha, R., Lin, T.J., Cai, Y., Cao, J., Zhang, Y., Li, H.: R^2-gaussian: Rectifying radiative gaussian splatting for tomographic reconstruction. In: Advances in Neural Information Processing Systems. vol. 37, pp. 44907–44934 (2024)