EUROGRAPHICS 2026 / B. Masia and J. Thies
(Guest Editors)

COMPUTER GRAPHICS *forum*
Volume 45 (2026), Number 2

Supplementary Material: Splat-based Metal Artifact Reduction in Cone-Beam CT via Polychromatic Modeling

Kiseok Choi Inchul Kim Jaemin Cho Hyeongjun Cho Min H. Kim

KAIST

1. Overview of Our Optimization Pipeline

We illustrate our overall optimization pipeline in Figure 1. Our method consists of two stages, following a structure similar to the baseline approach [ZLC*24]. In the first stage, a conventional reconstruction method, such as FDK [FDK84], is employed to obtain an initial volumetric reconstruction. From this volume, we randomly sample a subset of voxel locations and initialize 3D Gaussian primitives with their centers placed at the sampled positions. These Gaussians are forward-projected onto the projection image domain under a cone-beam geometry using the given system response parameters. The resulting synthetic projections are compared with the measured projection images, and the reconstruction loss is backpropagated to jointly optimize the corresponding 3D Gaussian parameters and the system response parameters. During this iterative optimization, our adaptive density control strategy is applied to prune the negligible Gaussians and to split or clone Gaussians in order to better represent the underlying volume. The output of the first stage is a set of optimized 3D Gaussian parameters. In the second stage, the optimized Gaussians are evaluated at the voxel locations of interest, from which the final reconstructed volume is obtained.

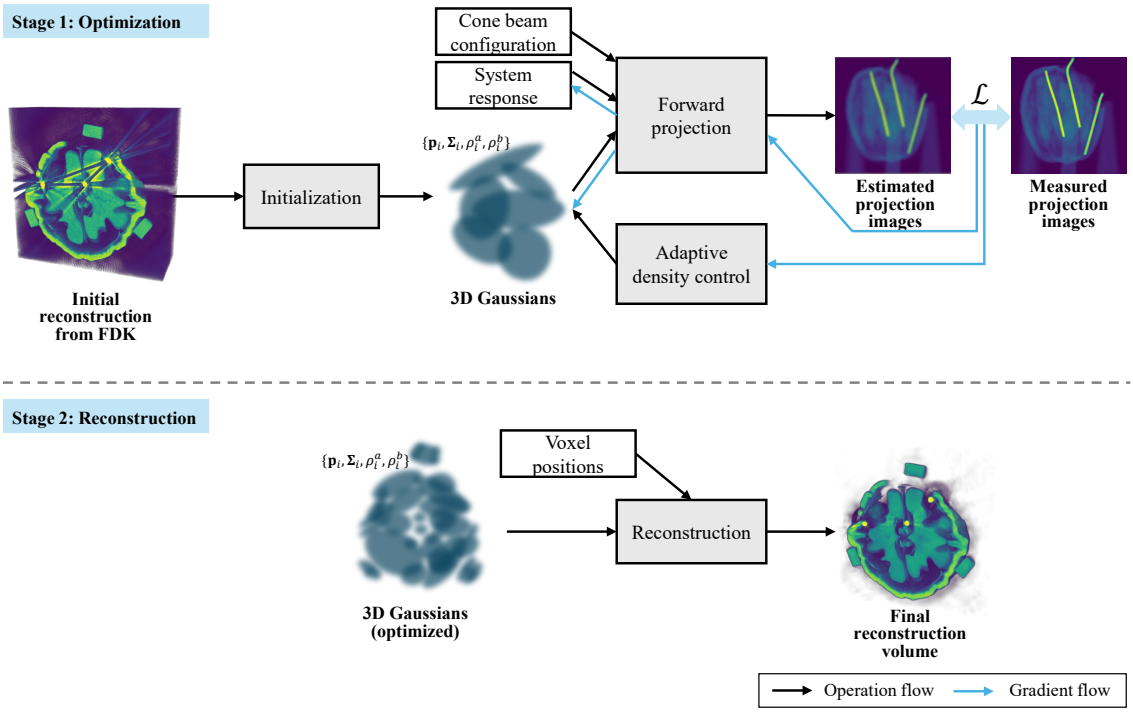

Figure 1: *Overview of our optimization pipeline.*

2. Additional Ablation Results

Figure 2 provides additional qualitative results that support the ablation experiments reported by Table 5 in the main paper. The numbers shown in the upper-right corners indicate the corresponding 3D PSNR values. As shown in this figure, the baseline reconstruction is noticeably blurred and suffers from severe metal artifacts. When only our response model is applied, finer details are recovered; however, prominent metal artifacts remain since the beam-hardening is not taken into account. Applying only the attenuation model significantly reduces these artifacts, although minor residual artifacts persist near metallic regions. When both models are jointly employed, metal artifacts are further suppressed and the overall intensity level closely matches the ground-truth volume, resulting in the highest PSNR among all ablation settings.

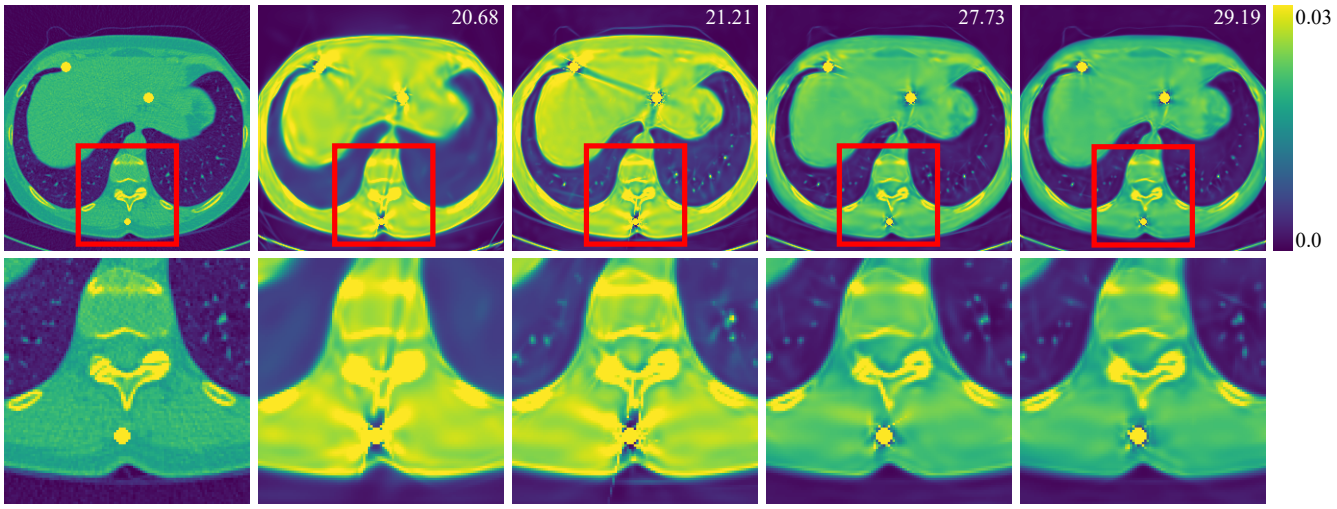

Figure 2: *Qualitative results of the ablations reported by Table 5 in the main paper.*

3. Derivation of Backward Propagation

We implement a custom CUDA pipeline to enable automatic differentiation of the parameters in the discrete polychromatic forward projection model using the Gaussian primitives as Equation (1):

$$P(\hat{\mathbf{x}}) = \sum_{i}^{M} s_i G_i(\hat{\mathbf{x}}) \rho_i^a - \log \sum_{k=0}^{N-1} \eta_k \exp\left(- \frac{1}{\left(\left(\frac{2k}{N-1} - 1 \right) \gamma + 1 \right)^3} \sum_{i}^{M} s_i G_i(\hat{\mathbf{x}}) \rho_i^b \right), \tag{1}$$

where $s_i = \sqrt{\frac{2\pi|\tilde{\Sigma}_i|}{|\Sigma_i|}}$, $G_i(\hat{\mathbf{x}}) = \exp\left(-\frac{1}{2} (\hat{\mathbf{x}} - \hat{\mathbf{p}}_i)^T \tilde{\Sigma}_i^{-1} (\hat{\mathbf{x}} - \hat{\mathbf{p}}_i) \right)$, $\sum_{k=0}^{N-1} \eta_k = 1$.

R^2-Gaussian [ZLC*24] consists of the `Preprocessor` that transforms the position and shape of each Gaussian, and the `Rasterizer` that computes the intensity projected onto each detector pixel. While the `Preprocessor` reuses the original implementation, we design a

new `Rasterizer` model and manually derive the gradients with respect to each intermediate parameter as Equations (2)–(6):

$$\left.\frac{d\mathcal{L}}{d\rho_i^a}\right|_{\hat{\mathbf{x}}} = \frac{d\mathcal{L}}{dP(\hat{\mathbf{x}})} \cdot \frac{dP(\hat{\mathbf{x}})}{d\rho_i^a} = \frac{d\mathcal{L}}{dP(\hat{\mathbf{x}})} \cdot (s_i G_i(\hat{\mathbf{x}})), \tag{2}$$

$$\left.\frac{d\mathcal{L}}{d\rho_i^b}\right|_{\hat{\mathbf{x}}} = \frac{d\mathcal{L}}{dP(\hat{\mathbf{x}})} \cdot \frac{dP(\hat{\mathbf{x}})}{d\rho_i^b} = \frac{d\mathcal{L}}{dP(\hat{\mathbf{x}})} \cdot \left(-\frac{1}{A(\hat{\mathbf{x}})} s_i G_i(\hat{\mathbf{x}}) \sum_{k=0}^{N-1} (B_k(\hat{\mathbf{x}}) C_k)\right), \tag{3}$$

$$\left.\frac{d\mathcal{L}}{ds_i}\right|_{\hat{\mathbf{x}}} = \frac{d\mathcal{L}}{dP(\hat{\mathbf{x}})} \cdot \frac{dP(\hat{\mathbf{x}})}{ds_i} = \frac{d\mathcal{L}}{dP(\hat{\mathbf{x}})} \cdot \left(\rho_i^a G_i(\hat{\mathbf{x}}) - \frac{1}{A(\hat{\mathbf{x}})} \rho_i^b G_i(\hat{\mathbf{x}}) \sum_{k=0}^{N-1} (B_k(\hat{\mathbf{x}}) C_k)\right), \tag{4}$$

$$\left.\frac{d\mathcal{L}}{dG_i}\right|_{\hat{\mathbf{x}}} = \frac{d\mathcal{L}}{dP(\hat{\mathbf{x}})} \cdot \frac{dP(\hat{\mathbf{x}})}{dG_i(\hat{\mathbf{x}})} = \frac{d\mathcal{L}}{dP(\hat{\mathbf{x}})} \cdot \left(\rho_i^a s_i - \frac{1}{A(\hat{\mathbf{x}})} \rho_i^b s_i \sum_{k=0}^{N-1} (B_k(\hat{\mathbf{x}}) C_k)\right), \tag{5}$$

$$\left.\frac{d\mathcal{L}}{d\eta_k}\right|_{\hat{\mathbf{x}}} = \frac{d\mathcal{L}}{dP(\hat{\mathbf{x}})} \cdot \frac{dP(\hat{\mathbf{x}})}{d\eta_k} = \frac{d\mathcal{L}}{dP(\hat{\mathbf{x}})} \cdot \left(-\frac{1}{A(\hat{\mathbf{x}})} \frac{B_k(\hat{\mathbf{x}})}{\eta_k}\right), \tag{6}$$

where $A(\hat{\mathbf{x}}) = \sum_{k=0}^{N-1} B_k(\hat{\mathbf{x}})$, $B_k(\hat{\mathbf{x}}) = \eta_k \exp\left(-C_k \sum_i^M \rho_i^b s_i G_i(\hat{\mathbf{x}})\right)$, $C_k = -\frac{1}{\left(\left(\frac{2k}{N-1}-1\right)\gamma+1\right)^3}$. All these equations are then implemented in the CUDA backward function. The final gradient with respect to η_k is obtained by summing the pixel-wise partial derivatives:

$$\frac{d\mathcal{L}}{d\eta_k} = \sum_{\hat{\mathbf{x}}} \left.\frac{d\mathcal{L}}{d\eta_k}\right|_{\hat{\mathbf{x}}} = \sum_{\hat{\mathbf{x}}} \frac{d\mathcal{L}}{dP(\hat{\mathbf{x}})} \cdot \left(-\frac{1}{A(\hat{\mathbf{x}})} \cdot \frac{B_k(\hat{\mathbf{x}})}{\eta_k}\right) \tag{7}$$

4. Comparison Implementation Detail

Since both Polyner [WCW*23] and Park et al. [PSJ] are originally designed for the fan-beam geometry, we extend their implementations to support the cone-beam configuration for fair comparison with our method. Specifically, we constrain the sampling range that is originally set to (-SOD,SOD) where SOD is Source-to-Object Distance, to a narrower interval that better matches the actual object region to reduce sensitivity to variations across different datasets. To improve the reconstruction quality under the cone-beam setting, which is more challenging to reconstruct a volume due to its 3D geometry of beams, we also increase the depth of the network and expand the hash encoding parameters used in the reconstruction. As the official implementation of Park et al. is not publicly available, we instead adapt the extended cone-beam implementation of Polyner and fine-tune several network and training-related parameters for fair comparison. For learning-based methods such as ACDNet, DICDNet, and OSCNet, we employ publicly available pre-trained models. As these baselines only support metal artifact reduction on 2D slices, we first reconstruct a 3D volume using FDK and then apply each model independently to the individual slices.

5. Geometry Configuration

The geometry configurations for the datasets are provided in Tables 1–2. To validate the robustness of our method, we conduct experiments under various geometric setups.

6. GPU Memory Usage

We report the memory footprint of the baseline and our method in Table 3. Although our model is more complex than the baseline, its memory requirement increases only marginally, indicating that the additional polychromatic features do not compromise memory efficiency.

7. Additional Discussion

We compute the PSNR and SSIM for the 3D volume to compare the reconstruction performance on the synthetic dataset. For the PSNR calculation, the ground truth (GT) is determined based on the center energy, and following the evaluation methods of existing metal artifact reduction techniques [WCW*23] and [PSJ], PSNR is measured for all regions except for the metal mask area. Since there is no proper GT available for the real dataset, we do not conduct quantitative evaluations, but qualitative comparison evaluations are done instead.

8. Additional Results

Figure 3 presents the objects for which we provide additional results. Figures 4–6 show qualitative comparisons of our method against FDK [FDK84] and the state-of-the-art joint reconstruction and metal artifact reduction methods: Polyner [WCW*23], and Park et al. [PSJ].

Table 1: *The scanning configuration of our synthetic datasets.*

Parameters	Synthetic Dataset		
	Broccoli	Lung	Teeth
Source voltage (kV)	90	←	←
Aluminium filter (mm)	0.5	←	←
Metal material	Aluminum	Steel	Titanium
Volume resolution	256×256×256	←	←
Volume size (mm)	50.0×50.0×50.0	←	←
Angle range (°)	[0, 360)	←	←
Number of angles	720	←	←
Detector resolution	512×512	←	←
Detector size (mm)	145.485	←	←
Source-to-object distance (mm)	200.962	←	←
Source-to-detector distance (mm)	501.309	←	←

Table 2: *The scanning configuration of our real datasets.*

Parameters	Real Dataset				
	Walnut	Metal Rods	Chicken (Wire)	Bell Pepper (Rivet)	Broccoli (Rivet)
Source voltage (kV)	90	←	←	←	←
Source current (μA)	166	←	300	←	←
Aluminum filter (mm)	1.0	0.5	1.0	←	←
Metal material	Steel	Aluminum	Steel	←	←
Volume resolution	256×256×256	←	←	←	←
Volume size (mm)	40.0×40.0×40.0	60.0×60.0×60.0	120.0×120.0×120.0	←	←
Angle range (°)	[0, 360)	←	←	←	←
Number of angles	720	←	←	←	←
Detector resolution	512×512	←	←	←	←
Detector size (mm)	104.773	145.485	229.902	145.485	229.902
Source-to-object distance (mm)	200.962	←	334.942	←	←
Source-to-detector distance (mm)	501.309	←	←	←	←

References

[FDK84] Feldkamp L. A., Davis L. C., Kress J. W.: Practical cone-beam algorithm. *Journal of the Optical Society of America A 1*, 6 (1984), 612–619. 1, 3

[PSJ] Park H. S., Seo J. K., Jeon K.: Implicit neural representation-based method for metal-induced beam hardening artifact reduction in x-ray ct imaging. *Medical Physics*. 3

[WCW*23] Wu Q., Chen L., Wang C., Wei H., Zhou S. K., Yu J., Zhang Y.: Unsupervised polychromatic neural representation for CT metal artifact reduction. In *Thirty-seventh Conference on Neural Information Processing Systems* (2023). 3

[ZLC*24] 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 (NeurIPS)* (2024). 1, 2, 5

Table 3: *GPU Memory usage (in MB) comparison between the baseline [ZLC*24] and our method across different scenes.*

Type	Scene	Baseline (MB)	Ours (MB)
Synthetic	Broccoli	1212	1402
	Lung	1242	1404
	Teeth	1272	1398
Real	Walnut	1308	1390
	Metal Rods	1330	1406
	Chicken (Wire)	1196	1382
	Bell Pepper (Rivet)	1210	1388
	Broccoli (Rivet)	1256	1340

Chicken (Wire) Bell Pepper (Rivet) Broccoli (Rivet)

Figure 3: *The objects used for our additional results. The type of metallic objects used for each scene is shown in the parentheses.*

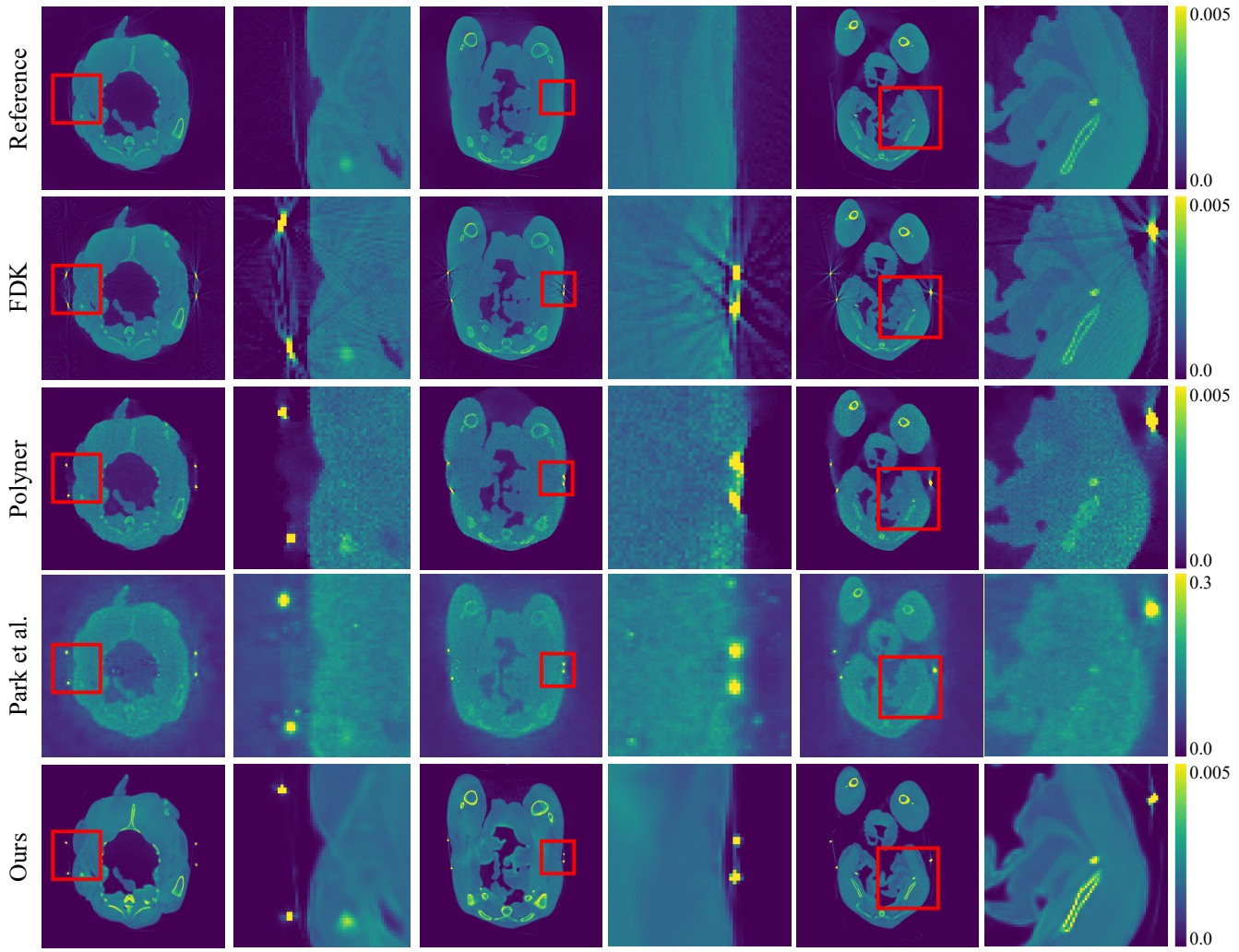

Figure 4: *Qualitative comparison on the real dataset: Chicken (Wire). The odd columns show the sampled slices from the reconstructed volumes by each algorithm, and the red boxes indicate the close-up regions. The even columns show the zoom-in patches from the corresponding slices.*

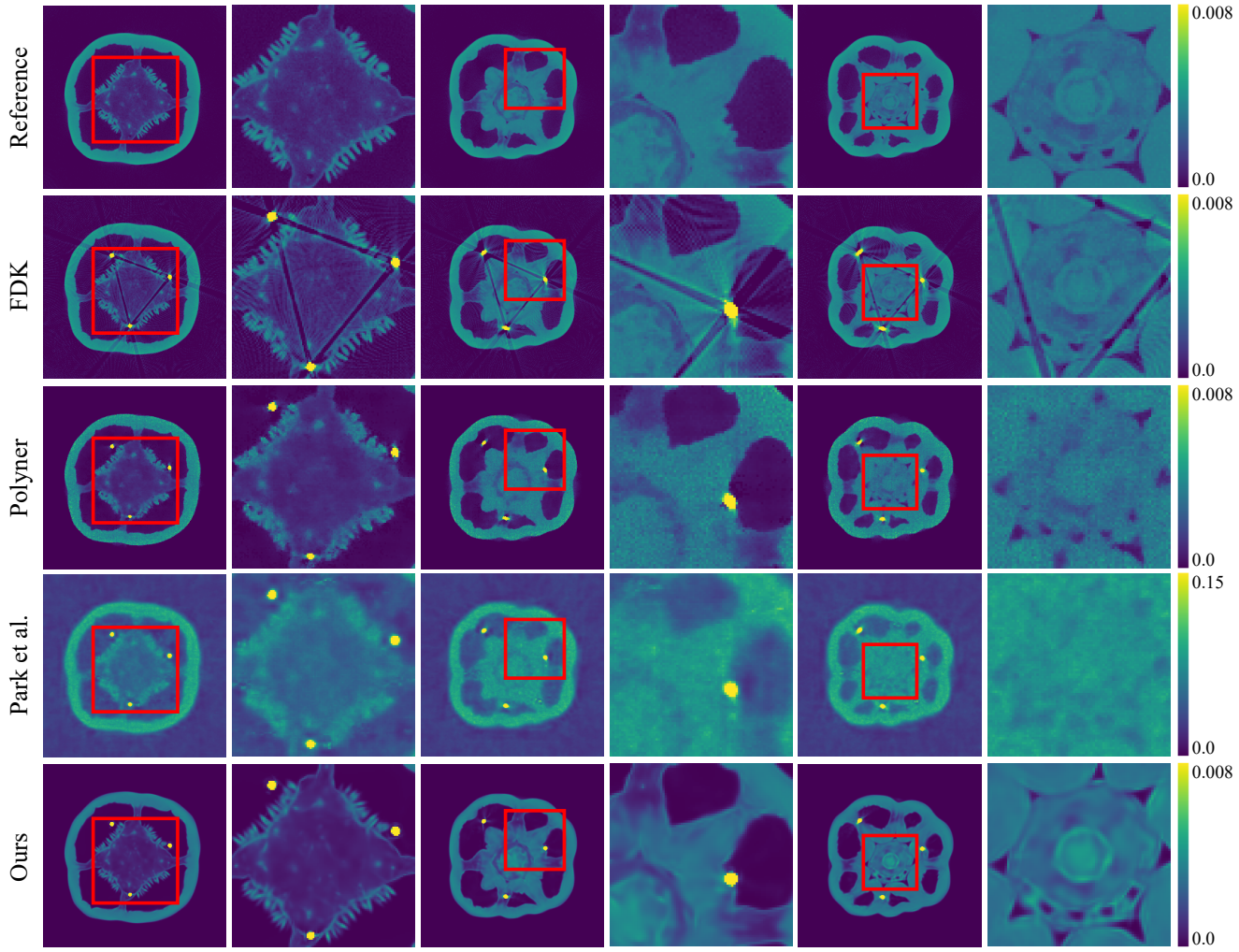

Figure 5: *Qualitative comparison on the real dataset: Bell Pepper (Rivet). The odd columns show the sampled slices from the reconstructed volumes by each algorithm, and the red boxes indicate the close-up regions. The even columns show the zoom-in patches from the corresponding slices.*

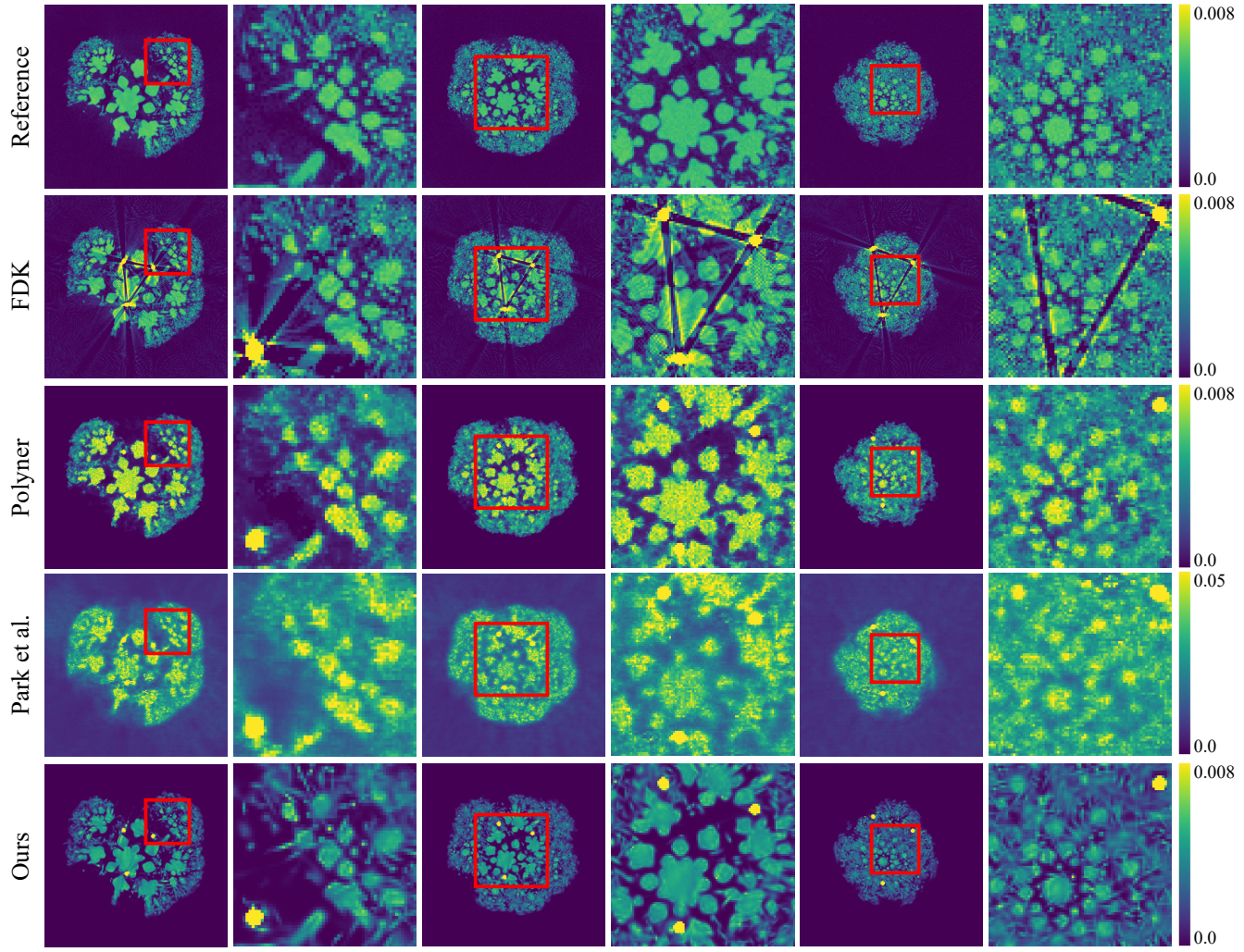

Figure 6: *Qualitative comparison on the real dataset: Broccoli (Rivet). The odd columns show the sampled slices from the reconstructed volumes by each algorithm, and the red boxes indicate the close-up regions. The even columns show the zoom-in patches from the corresponding slices.*