X-Field: A Physically Informed Representation for 3D X-ray Reconstruction

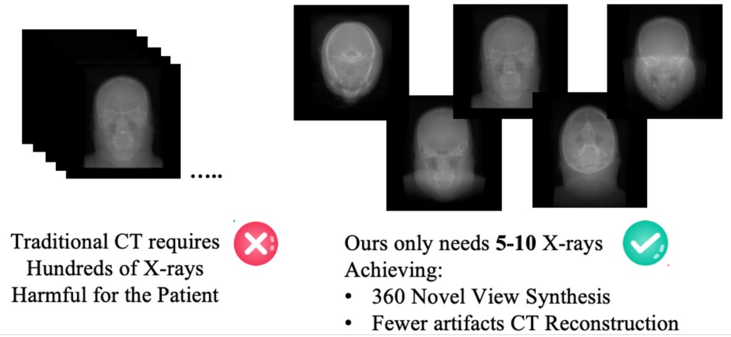
Feiran Wang[1*] **Jiachen Tao**[1*] **Junyi Wu**[1*] **Haoxuan Wang**[1] **Bin Duan**[2]
Kai Wang[3] **Zongxin Yang**[4] **Yan Yan**[1†]

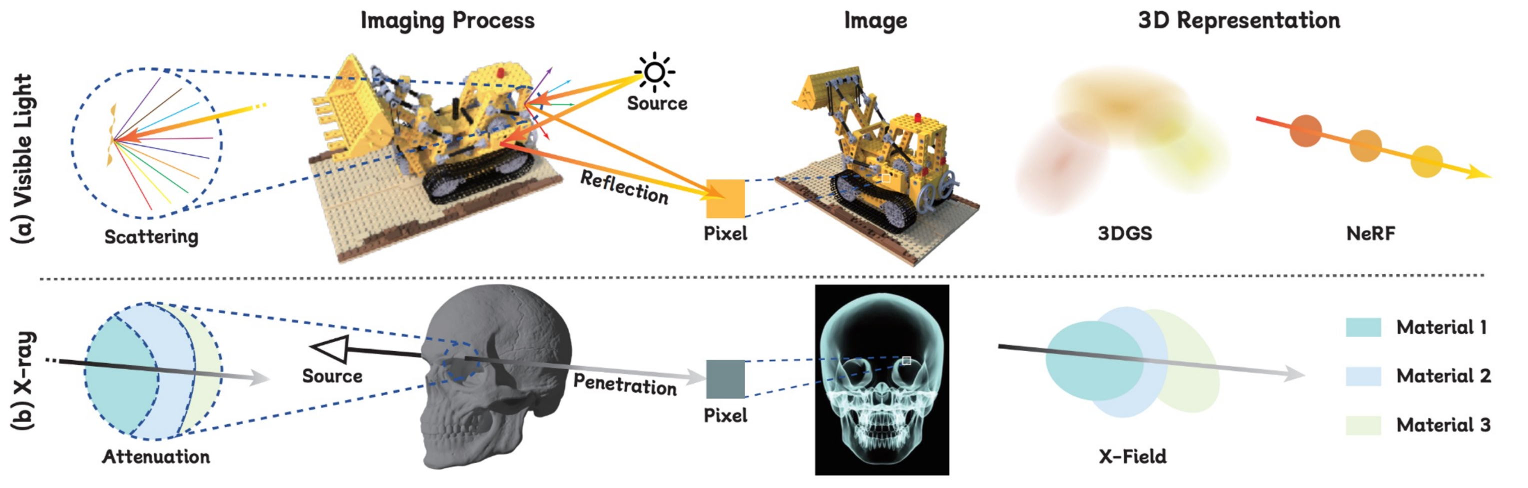
[1]University of Illinois Chicago [2]University of Michigan
[3]National University of Singapore [4]Harvard Medical School

Introduction

Task Definition

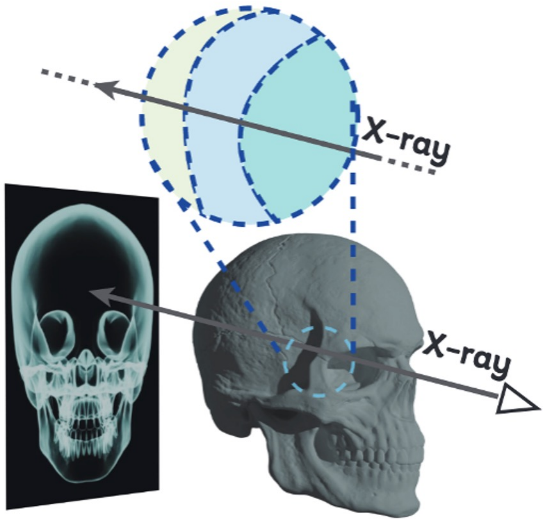
Following the X-ray imaging principle, introducing a Physically Grounded Representation for achieving X-ray Novel View Synthesis and CT Reconstruction with Highly Sparse X-ray input.





Imaging Principle: Visible Light VS X-ray

Physically Informed Ellipsoid Representation



let $\mathbf{r}(t) = \mathbf{o} + t\mathbf{d} \in \mathbb{R}^3$ denote an X-ray path,

$$I(\mathbf{r}) = \log I_0 - \log I'(\mathbf{r}) = \int_{t_0}^{t_n} \sigma(\mathbf{r}(t))\, dt.$$

$$\begin{aligned} I(\mathbf{r}) = \int_{t_0}^{t_n} \sigma(\mathbf{r}(t))\, dt &= \int_{t_0}^{t_1} \sigma(\mathbf{r}(t))\, dt + \cdots + \int_{t_{n-1}}^{t_n} \sigma(\mathbf{r}(t))\, dt \\ &= \sigma_0 \int_{t_0}^{t_1} dt + \cdots + \sigma_{n-1} \int_{t_{n-1}}^{t_n} dt \\ &= \sigma_0 l_0 + \sigma_1 l_1 + \cdots + \sigma_{n-1} l_{n-1}, \end{aligned}$$

Each ellipsoid has :

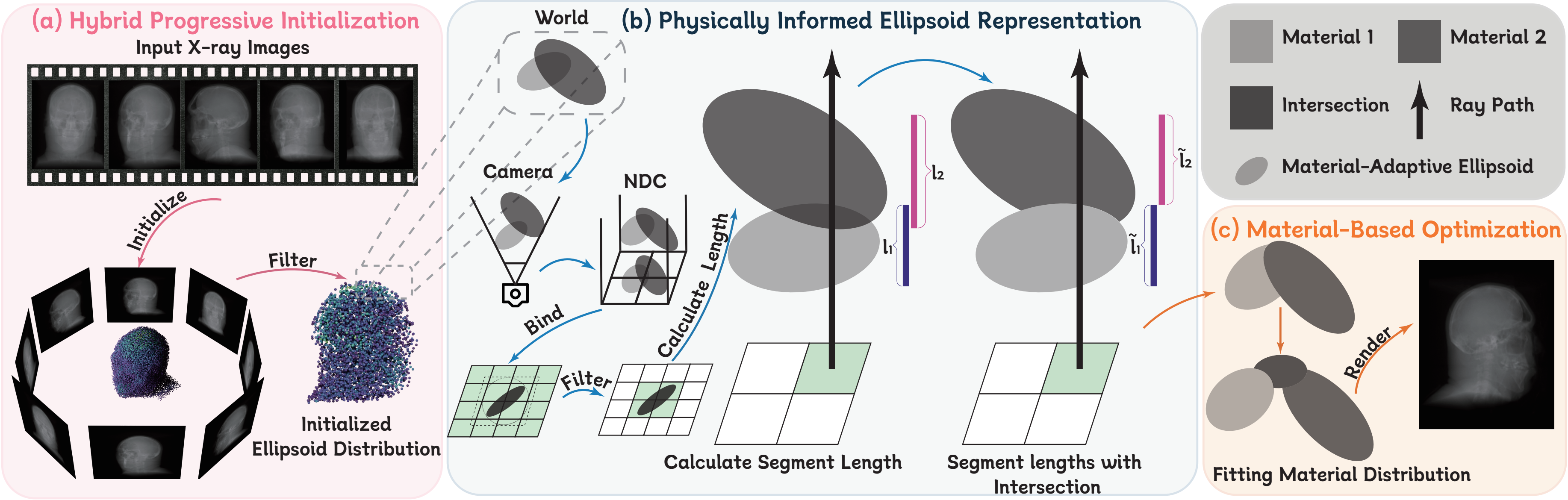
- a constant attenuation coefficient
- a defined segment length

$$l_i = l_{\max} \times \sqrt{1 - \left(\frac{C - B^2}{A}\right)}, \text{ where}$$

$$A = \mathbf{d}^\top \mathbf{\Sigma}_{3D}^{-1} \mathbf{d}, \quad B = \mathbf{a}^\top \mathbf{\Sigma}_{3D}^{-1} \mathbf{d}, \quad C = \mathbf{a}^\top \mathbf{\Sigma}_{3D}^{-1} \mathbf{a}.$$

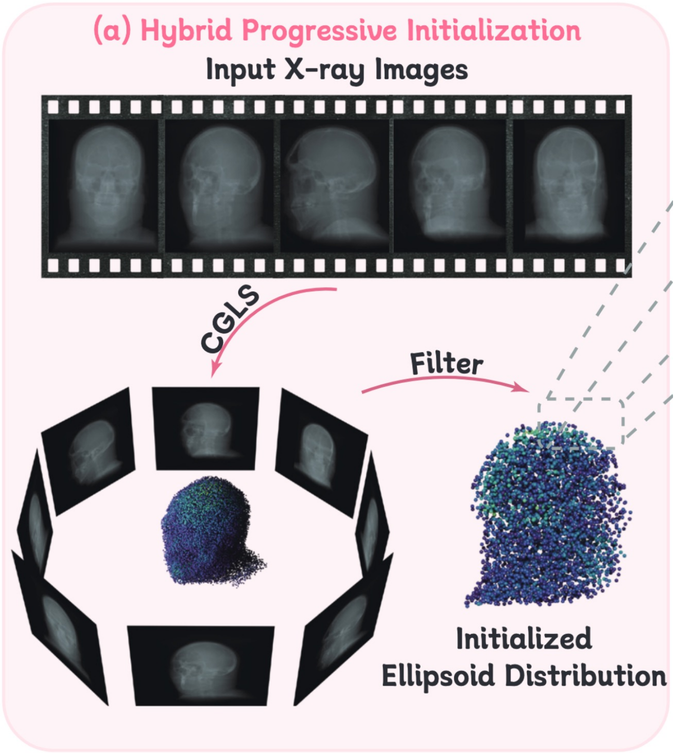
Pipeline Overview

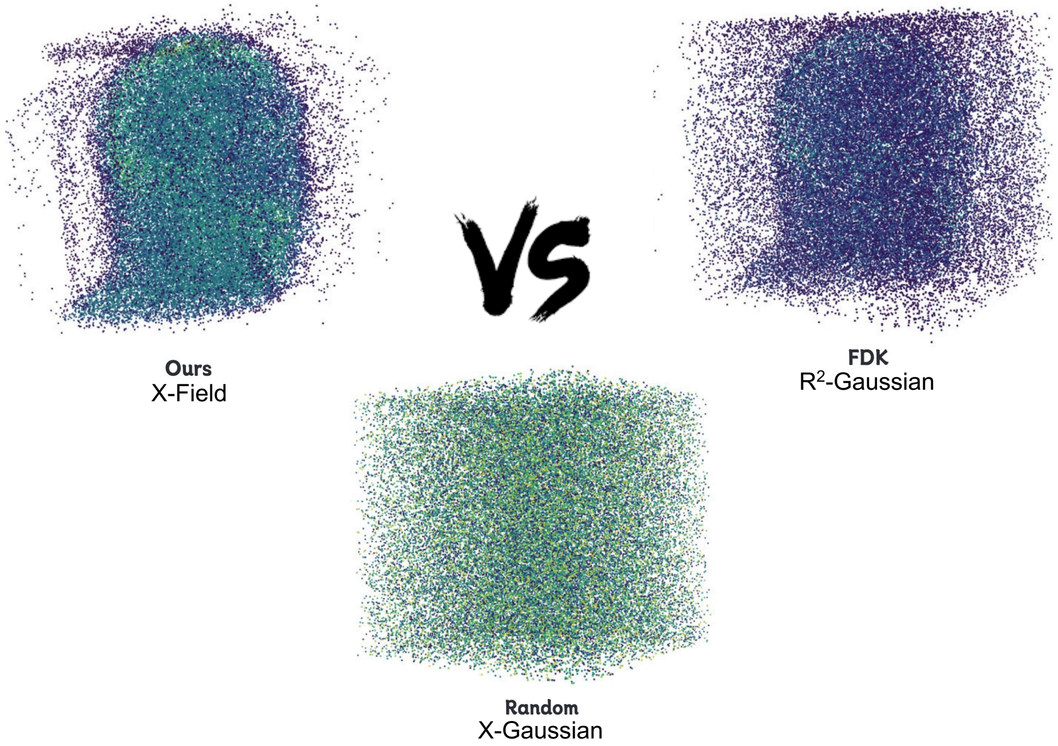
Our pipeline consists of three stages: initialization, physically informed X-ray field construction, and material-based optimization.



Hybrid Progressive Initialization

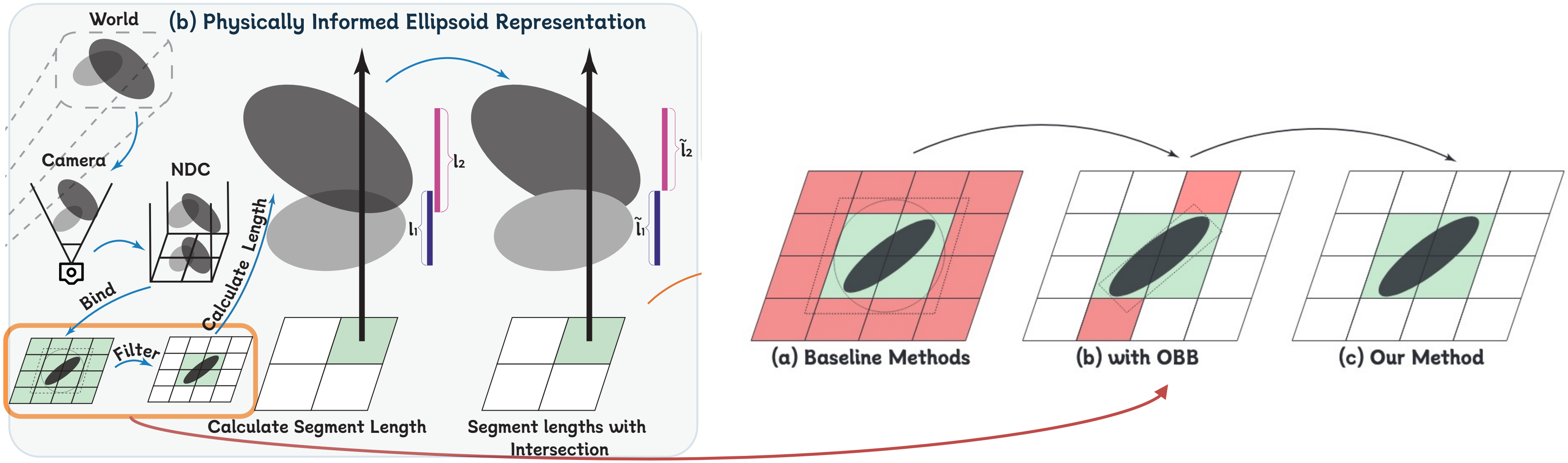
We introduce a hybrid progressive initialization that progressively refines ellipsoid distribution from coarse to fine, providing better geometric priors for faster and more stable optimization.





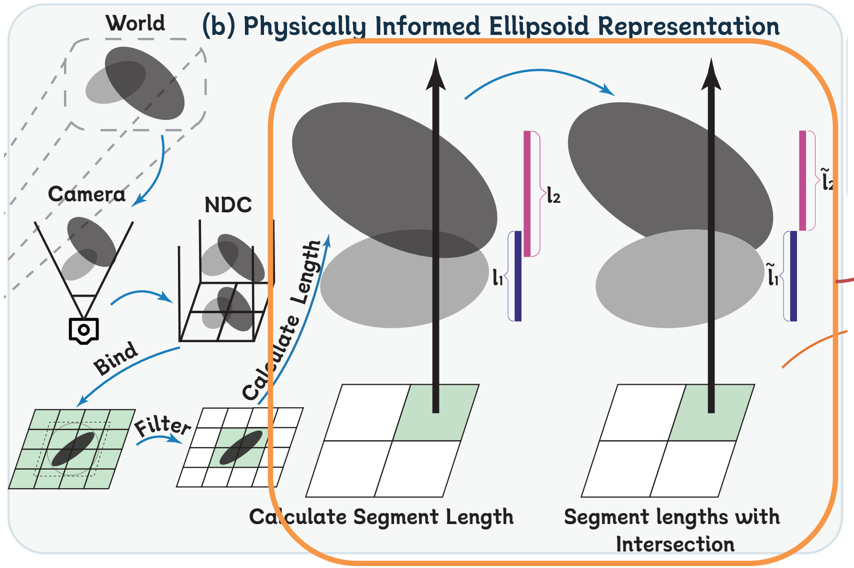
Physically Faithful Overlap Filtering

We filter pixel–ellipsoid overlaps based on physical feasibility, replacing AABB association with adaptive OBB mapping to avoid redundant computations.



Segment Length Computation with Intersections

We design an intersection-aware algorithm to ensure each spatial region is assigned to only one material, avoiding double counting of attenuation.



Algorithm 1 Compute Segment Lengths with Intersections

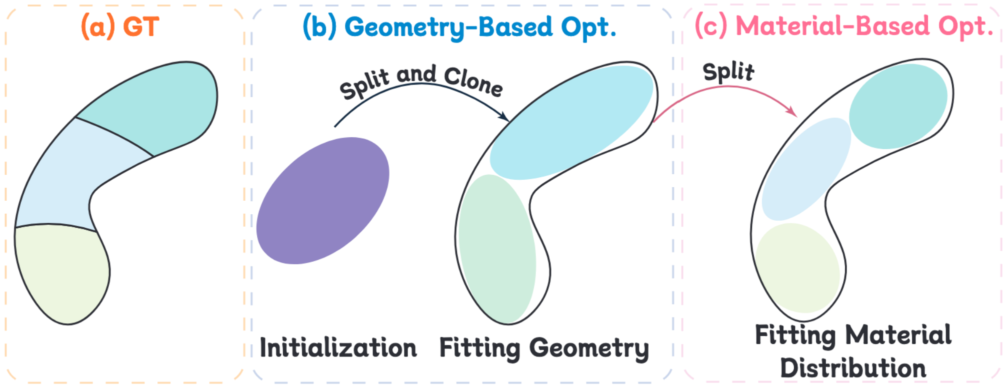
Input: $(z_0, z_1, \dots, z_{n-1})$: sorted depths of ellipsoids $\{\mathbf{E}_i\}$
$(l_0, l_1, \dots, l_{n-1})$: segment lengths for individual ellipsoids without considering intersections

Output: Updated segment lengths $\tilde{l}_0, \tilde{l}_1, \dots, \tilde{l}_{n-1}$ and effective regions

```
for i = 0 to n − 1 do
    if i == 0 then
        l̃_0 ← l_0
        z ← z_0, l ← l_0
    else
        if z_i < z + ½ l then
            l̃_i ← max(0, (z_i + ½ l_i) − (z + ½ l))
        else
            l̃_i ← min(l_i, (½ l_i + z_i) − (z + ½ l))
        end if
        if l̃_i ≠ 0 then
            Update the valid region of ellipsoid E_i as [z_i + ½ l_i − l̃_i, z_i + ½ l_i]
            z ← z_i, l ← l_i
        end if
    end if
end for
```

Material-Based Optimization

Beyond fitting geometry, our material-based optimization adaptively refines ellipsoids near material boundaries to capture attenuation differences between tissues.



Method	PSNR ↑	SSIM ↑	LPIPS* ↓
w/o Material Opt.	34.78	0.941	73.45
w/o Overlap Filter	34.59	0.937	74.32
w/o Intersection	33.84	0.929	76.60
w/o Ray Length	27.48	0.875	87.83
Ours	35.03	0.953	72.12

Table 2. **Ablation on the Proposed Components (§ 5.3).**

Quantitative Results – Novel View Synthesis

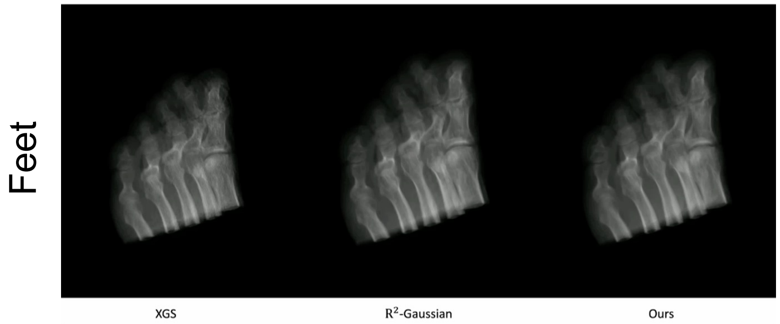
Our X-Field achieves the best overall performance on both real and synthetic datasets, outperforming R^2-Gaussian across most metrics and input settings.

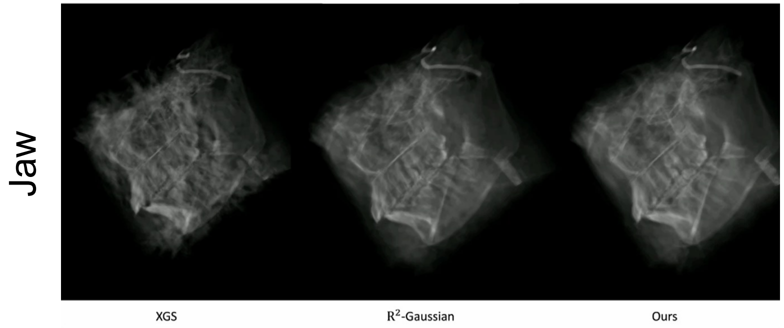
Method	Human Organ 10-views [16]			Daily Object 10-views [16]			Human Organ 5-views [16]			Daily Object 5-views [16]		
	PSNR↑	SSIM↑	LPIPS*↓	PSNR↑	SSIM↑	LPIPS*↓	PSNR↑	SSIM↑	LPIPS*↓	PSNR↑	SSIM↑	LPIPS*↓
Traditional Methods												
FDK [3]	12.35	0.675	291.2	16.52	0.716	259.1	8.15	0.618	310.6	14.42	0.688	283.7
SART [4]	13.23	0.691	284.8	17.69	0.724	247.3	9.31	0.634	303.4	15.68	0.663	293.5
Deep Learning-based Methods												
TensoRF [12]	16.61	0.928	182.5	24.19	0.946	153.4	12.32	0.895	189.6	18.27	0.922	210.8
NeAT [17]	16.22	0.934	185.3	25.15	0.957	155.2	11.08	0.887	188.3	17.29	0.908	211.3
NAF [13]	17.89	0.925	193.2	25.44	0.949	151.9	11.19	0.894	197.1	17.02	0.923	208.5
SAX-NeRF [14]	19.32	0.945	186.4	25.38	0.979	143.6	14.18	0.901	191.2	19.09	0.948	204.9
X-Gaussian [15]	22.88	0.947	130.3	22.91	0.982	79.12	17.23	0.947	176.4	20.31	0.961	108.1
R^2-Gaussian [16]	33.72	0.967	85.97	41.93	0.986	54.31	31.12	0.956	109.7	34.52	0.965	82.46
Ours	35.71	0.980	71.03	42.80	0.983	45.64	32.34	0.963	103.2	37.41	0.970	81.02

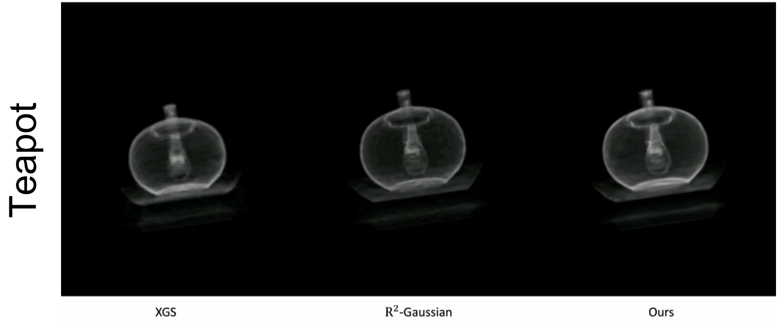
Table 1: **Quantitative Comparison of NVS.** We compare our X-Field with: (a) Traditional analytical method: FDK , SART . (b) Deep Learning-based methods: TensoRF , NeAT , NAF , SAX-NeRF , X-Gaussian , and R^2-Gaussian . We report LPIPS* = LPIPS $\times 10^3$. We mark out best and second best method for all metrics.

Qualititive Results – Novel View Synthesis

Our method synthesizes sharper, artifact-free X-ray views across a full 360° range, preserving fine anatomical and structural details.

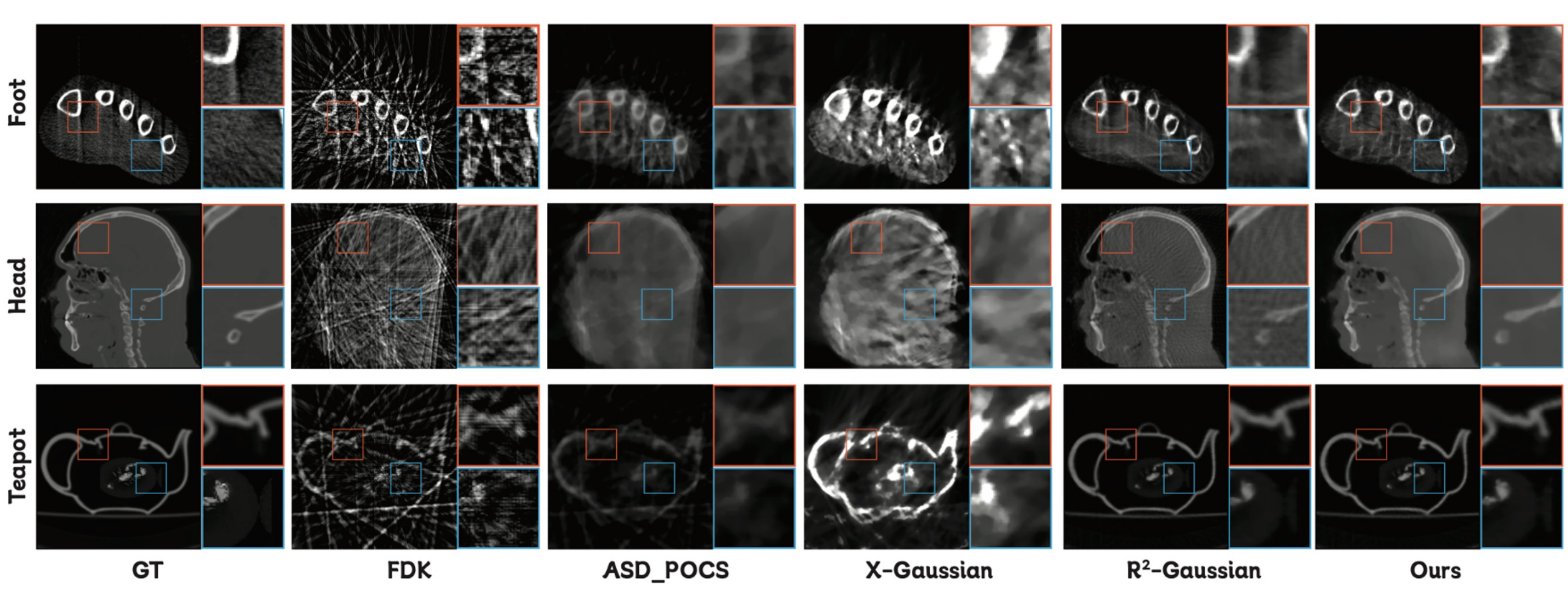






Qualititive Results – CT Reconstruction

Our method reconstructs CT volumes with clearer textures, sharper anatomical boundaries, and fewer artifacts, especially in high-contrast regions such as bone structures and cranial cavities.



Thanks