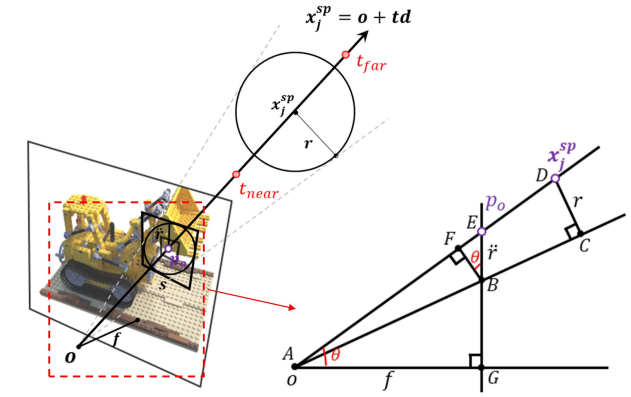
HashPoint: Accelerated Point Searching and Sampling for Neural Rendering

Supplementary Material



(A) Adaptive searching kernel (B) Derivation diagram

Figure 10. (A) Adaptive searching kernel applied on an image plane. The square represents a pixel to be rendered on the image. The circle represents a cross-section of the searching kernel of a cone to select points to render. (B) Detailed derivation of the searching kernel.

In this supplementary material, we provide details about the following topics:

- *Derivation of searching radius* in Appendix 1;
- *Additional experiments* in Appendix 2;
- *Breakdown results* in Appendix 3;
- *Visualization* in Appendix 4;
- *Limitation* in Appendix 5.

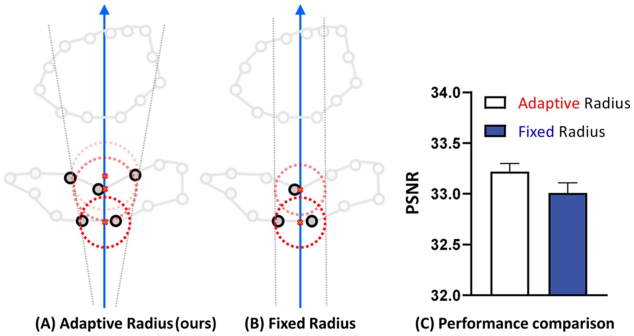
1. Adaptive searching radius derivations

In this section, we elaborate on the derivation of adaptive searching kernel r as mentioned in Section 3.2. We initiate the derivation with a disc representation on the image plane, approximating the area of the pixel. The radius of the disc can be calculated by $\dot{r} = \sqrt{\Delta x \cdot \Delta y / \pi}$, where Δx and Δ_y are the width and height of the pixel in world coordinates. For a broader coverage area, we shift from a disc with radius $\dot{r}$ for one pixel to a larger disc with radius $\ddot{r}$ for the searching kernel. The searching kernel size s can be expressed as $s = 2 \cdot \lceil \frac{\ddot{r}}{\dot{r}} \rceil + 1$. Specifically, given the searching kernel s and radius $\dot{r}$, the suitable range for $\ddot{r}$ is,

$$\left(\frac{s-1}{2} - 1\right) \cdot \dot{r} < \ddot{r} \leq \left(\frac{s-1}{2}\right) \cdot \dot{r}. \tag{6}$$

As shown in Figure 10B, to determine the radius r for sample points, we employ the concept of similar triangles, specifically $\triangle ABF \sim \triangle ADC$. This set of similar triangles establishes a proportionality that allows us to solve for r as,

$$r = CD = \frac{AD \times BF}{AB}. \tag{7}$$



(A) Adaptive Radius (ours) (B) Fixed Radius (C) Performance comparison

Figure 11. Radius Strategies and Performance: (A) depicts an adaptive radius correlated with viewport distance, (B) shows a constant fixed radius, and (C) compares their PSNR performance. Additional comparison of point cloud selection strategies in Figure 4 in the main text.

Here, AD represents the Euclidean distance between the camera ray $\mathbf{o}$ and the sample point position $\mathbf{x_j^{sp}}$,given by $AD = \left\| \mathbf{x_j^{sp}} - \mathbf{o} \right\|_2$. By invoking another set of similar triangles, $\triangle BEF \sim \triangle AEG$, we can calculate BF as,

$$BF = \ddot{r} \cdot cos\theta = \frac{\ddot{r} \cdot AG}{AE} = \frac{\ddot{r} f}{\|\mathbf{p_o} - \mathbf{o}\|_2}, \tag{8}$$

where f is the local length and $\mathbf{p_o}$ is the pixel center intersected by the camera ray. The length AB is the hypotenuse of the right triangle $\triangle AGB$ and is computed using the Pythagorean theorem as,

$$\begin{aligned} AB &= \sqrt{GB^2 + AG^2} \\ &= \sqrt{(GE - BE)^2 + AG^2} \\ &= \sqrt{\left(\sqrt{AE^2 - AG^2} - BE\right)^2 + AG^2} \\ &= \sqrt{\left(\sqrt{\|\mathbf{p_o} - \mathbf{o}\|_2^2 - f^2} - \ddot{r}\right)^2 + f^2} \end{aligned} \tag{9}$$

Combining these elements, the final expression for the radius r of the searching kernel is,

$$r = \frac{\|\mathbf{x_j^{sp}} - \mathbf{o}\|_2 \cdot f\ddot{r}}{\|\mathbf{p_o} - \mathbf{o}\|_2 \cdot \sqrt{\left(\sqrt{\|\mathbf{p_o} - \mathbf{o}\|_2^2 - f^2} - \ddot{r}\right)^2 + f^2}}. \tag{10}$$

This integrated expression encapsulates the geometric relationship between the image plane, the searching kernel,

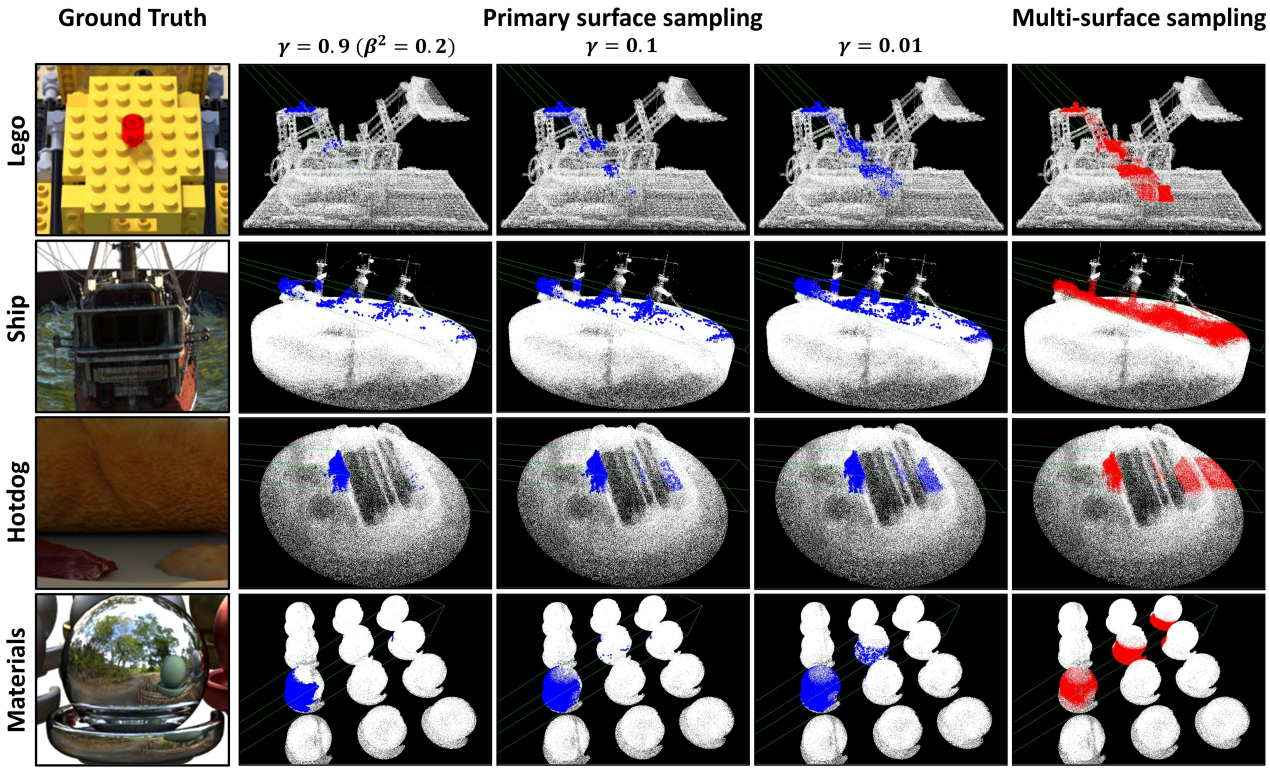


Figure 12. Adaptive sampling for noisy input, showing the transition from primary-surface to multiple-surface sampling by adjusting γ values, with visual results for different surfaces. It features a visual comparison across different surfaces. The first column displays the ground truth, and the last column showcases the region where multi-surface sampling is applied in Point-NeRF. The middle part illustrates the gradual transition. In the main text, the default value of β^2 is set to 0.02 when dealing with refined geometry input.

and the camera parameters, completing the derivation of the adaptive searching kernel's radius.

Room for improvement. As the angle $\widehat{BOE}$ is minuscule (less than 1 degree), we can make the below approximation and then apply a similar triangle relationship:

$$AB \approx AF = AE - EF = AE - \frac{BF \times EG}{AG} \tag{11}$$

In fact, we can further simplify eq. 11 by applying small angle assumption of $\widehat{BOE}$ twice: $AB \approx AF \approx AE$. Eq. 7 becomes:

$$\begin{aligned} r &\approx \frac{AD \times BF}{AE} = \frac{\|\mathbf{x_j^{sp}} - \mathbf{o}\|_2 \cdot \frac{\ddot{r}f}{\|\mathbf{p_o} - \mathbf{o}\|_2}}{\|\mathbf{p_o} - \mathbf{o}\|_2} \\ &\approx \frac{\|\mathbf{x_j^{sp}} - \mathbf{o}\|_2 \cdot \ddot{r}f}{\|\mathbf{p_o} - \mathbf{o}\|_2^2} \end{aligned} \tag{12}$$

Eq. 12 is simpler and faster to computer than eq. 10.

2. Additional experiments

Comparison between the fixed radius and our adaptive radius sampling. Regarding the discussion in Section 3, we compared the performance differences between adaptive radius and fixed radius on the NeRF-Synthesis dataset, building upon the Point-NeRF framework. Figure 11 illustrates the differences between adaptive and fixed radius strategies for point cloud sampling in neural radiance fields. As explained in Section 3.5, the adaptive radius expands with increased distance from the ray origin, resembling a camera's view frustum, while the fixed radius used by Point-NeRF remains constant. The performance in Figure 11C demonstrates that the adaptive radius yields higher PSNR values, indicating a reduction in noise by selecting only points relevant to the rendered pixel.

Our adaptive sampling for noisy input. This section extends the discussion on the solution to noisy input from Section 4.3. Primary surface sampling encounters challenges in accurately extracting surface features from noisy point cloud input. As outlined in the ablation study, we proposed two methods to address issues with noisy inputs. The first involves a brief 10-minute geometry optimization that significantly enhances the benefit of our methods by filling gaps and refining noisy point clouds from noisy surfaces. The second method adaptively adjusts the sampling scope, shifting from primary surface to multi-surface sampling. Notably, unlike full multi-surface sampling, which gathers features from all intersected surfaces, our approach, as illustrated in Figure 12, samples just slightly beyond the primary surface, avoiding extensive multi-surface feature collection.

Method	Rm 0	Rm 1	Rm 2	Off 0	Off 1	Off 2	Off 3	Off 4	Avg.
NICE-SLAM [69]	0.97	1.31	1.07	0.88	1.00	1.06	1.10	1.13	1.06
Point-SLAM [48]	**0.61**	**0.41**	**0.37**	**0.38**	0.48	**0.54**	0.69	0.72	**0.52**
Point-SLAM + Ours	0.69	0.53	**0.37**	0.47	**0.45**	0.65	**0.64**	**0.54**	0.54

Table 6. Tracking performance on Replica [53] (ATE RMSE ↓ [cm]).

Method	Metric	Rm 0	Rm 1	Rm 2	Off 0	Off 1	Off 2	Off 3	Off 4	Avg.
NICE-SLAM[69]	Depth L1 [cm] ↓	1.81	1.44	2.04	1.39	1.76	8.33	2.01	2.97	2.97
	Precision [%] ↑	45.86	43.76	44.38	51.40	50.80	38.37	40.85	37.35	44.10
	Recall [%] ↑	44.10	46.12	42.78	48.66	53.08	39.98	39.24	35.77	43.69
	F1 [%] ↑	44.96	44.84	43.56	49.99	51.91	39.16	39.92	36.54	43.86
Point-SLAM[48]	Depth L1 [cm] ↓	0.53	**0.22**	**0.46**	**0.30**	0.57	**0.49**	**0.51**	0.46	**0.44**
	Precision [%] ↑	91.95	**99.04**	**97.89**	**99.00**	**99.37**	98.05	96.61	**93.98**	96.99
	Recall [%] ↑	82.48	**86.43**	84.64	**89.06**	84.99	81.44	81.17	78.51	83.59
	F1 [%] ↑	86.90	**92.31**	**90.78**	**93.77**	91.62	88.98	88.22	85.55	89.77
Point-SLAM + Ours	Depth L1 [cm] ↓	**0.51**	0.23	0.55	0.37	**0.46**	0.54	0.52	**0.42**	0.45
	Precision [%] ↑	**92.31**	98.96	97.56	98.85	99.31	**99.33**	**97.45**	93.95	**97.22**
	Recall [%] ↑	**83.32**	85.98	**84.65**	88.86	**85.23**	**81.74**	**82.71**	**79.38**	**83.98**
	F1 [%] ↑	**87.58**	92.01	90.64	93.58	**91.73**	**89.68**	**89.38**	**86.05**	**90.11**

Table 7. Reconstruction performance on Replica [53].

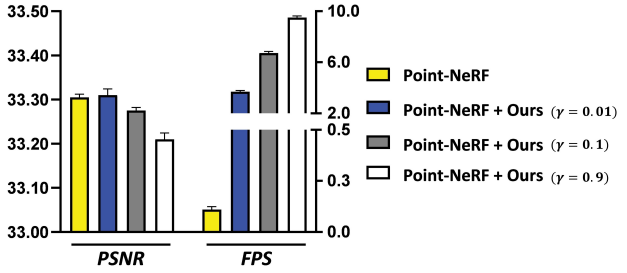


Figure 13. The illustration of the trade-off between efficiency and accuracy in the transition from primary surface to multiple surface sampling under the noisy input by adjusting the γ parameter. Point-NeRF with our HashPoint and $\gamma = 0.01$ outperforms original Point-NeRF in both *PSNR* and *FPS*.

Figure 13 demonstrates that reducing γ increases precision and reduces the speed of sampling. A larger γ (*e.g.*, 0.9) focuses on primary surface sampling, maximizing speed for interactive frame rate while maintaining competitive precision. Conversely, a smaller γ captures more point cloud features below the primary surface, exceeding the precision of multi-surface sampling and still achieving speeds over 20 times faster. In summary, the γ adjustment effectively balances sampling speed and precision, with each setting offering unique benefits in efficiency and accuracy.

Tracking and mapping. We supplement the integration of our method with Point-SLAM [48] for tracking and reconstruction on the Replica [53] dataset. As shown in Tables 6 and 7, our approach outperforms the original depth-guided sampling in terms of precision and recall, while maintaining competitive performance in other metrics.

Method		Performance	
Searching	Selection	PSNR ↑	FPS ↑
Point-NeRF			
Uniform grid	Multiple surface	33.31	0.12
HashPoint	Multiple surface	33.34	0.78
Uniform grid	Primary surface	33.20	6.92
HashPoint	Primary surface	33.22	9.60
Point-SLAM			
Depth	Single surface	30.01	4.64
Depth	Primary surface	30.32	6.72
Pointersect			
Uniform grid	K nearest points	31.23	1.30
HashPoint	K nearest points	31.23	1.75
Uniform grid	Primary surface	30.01	8.89
HashPoint	Primary surface	30.03	9.13
NPLF			
Brute force	K nearest points	30.44	0.48
HashPoint	K nearest points	30.42	3.12
Brute force	Primary surface	29.88	0.50
HashPoint	Primary surface	29.90	3.40

Table 8. Ablation study evaluates on NeRF-Synthesis dataset.

3. Breakdown results

Table 9 provides detailed per-scene quantitative results comparing integration with Point-NeRF on the NeRF-Synthesis dataset. Our integration significantly accelerates the process while preserving competitive performance. The comparative data for our ablation study is outlined in Table 8.

	Chair	Drums	Lego	Mic	Materials	Ship	Hotdog	Ficus
				PSNR↑				
NPBG	26.47	21.53	24.84	26.62	21.58	21.83	29.01	24.60
NeRF	33.00	25.01	32.54	32.91	29.62	28.65	36.18	30.13
NSVF	33.19	25.18	32.54	34.21	29.62	27.93	37.14	31.23
Point-NeRF	35.40	26.06	**35.04**	35.95	29.61	30.97	**37.30**	**36.13**
Point-NeRF + Ours	**35.54**	**26.12**	34.68	**36.34**	**30.64**	**31.08**	37.02	34.30
				SSIM↑				
NPBG	0.939	0.904	0.923	0.959	0.887	0.866	0.964	0.940
NeRF	0.967	0.925	0.961	0.980	0.949	0.856	0.974	0.964
NSVF	0.968	0.931	0.960	**0.987**	**0.973**	0.854	0.980	0.973
Point-NeRF	**0.984**	**0.935**	**0.978**	**0.990**	0.948	0.892	**0.982**	0.987
Point-NeRF + Ours	0.977	0.931	0.967	0.984	0.949	**0.920**	0.978	0.978
				$\text{LPIPS}_{vgg}\downarrow$				
NPBG	0.085	0.112	0.119	0.060	0.134	0.210	0.075	**0.078**
NeRF	0.046	0.091	0.050	0.028	**0.063**	0.206	0.121	0.044
Point-NeRF	**0.023**	**0.078**	**0.024**	**0.014**	0.072	0.124	0.037	**0.022**
Point-NeRF + Ours	0.028	0.101	0.047	0.018	0.075	**0.097**	**0.036**	0.041
				$\text{LPIPS}_{alex}\downarrow$				
NSVF	0.043	0.069	0.029	0.010	**0.021**	0.162	0.025	0.017
Point-NeRF	**0.010**	**0.055**	**0.011**	**0.007**	0.041	0.070	**0.016**	**0.009**
Point-NeRF + Ours	0.012	0.067	0.014	0.010	0.049	**0.054**	0.018	0.017

Table 9. Quantitative results in the NeRF Synthetic dataset.

Figure 14. The qualitative results of Point-SLAM [62] with ours on the Replica [53] dataset.

4. Visualization

We also experiment with the integration with Point-SLAM and NPLF on Replica and Waymo datasets respectively. The qualitative results are shown in Figure 14 and 15. Please find more visual results for sampling comparison in our video.

5. Limitation

During optimization, due to gradient propagation issues, multi-surface sampling is still necessary to sample and optimize all points as much as possible. Currently, our choice of β is fixed and does not dynamically adjust based on the geometry's distribution and noise level. Future work could ex-

Figure 15. The qualitative results of NPLF [41] with ours on the Waymo [54] dataset.

plore dynamic adjustment of the sampling process. For instance, as the geometry is progressively optimized, β could increase gradually, transitioning from multi-surface to primary surface sampling.