Hermite Radial Basis Function for Surface Reconstruction via Differentiable Rendering

Supplementary Material

This supplementary document provides additional details and results that complement the main paper. In Section 1, we provide the detailed mathematical proof of Theorem 1 and justify the choice of the HRBF polynomial offset $p_0 > 0$. Section 2 derives the closed-form initialization of the HRBF weighting parameters used in our implicit function. Section 3 analyzes the proposed radiance field parameterization in more depth, presenting quantitative PSNR results on DTU and BlendedMVS, studying the impact of using spherical harmonics and spherical Gaussians for radiance representation, and examining the coupling between the radiance field and the underlying implicit geometry. Sections 4 and 5 detail the evaluation protocols used on the DTU dataset [1] and the BlendedMVS dataset [11], respectively, including data preprocessing. Section 6 presents additional qualitative results on DTU [1] and BlendedMVS [11], illustrating the robustness and visual quality of our reconstruction. Finally, Section 7 provides additional quantitative results on the Tanks and Temples dataset.

1. Choice of the HRBF Polynomial Offset: Proof of Theorem 1

We analyze the volume rendering equation in the limit where the sharpness parameter s in the density model from Objects as Volumes [7] tends to $+\infty$. Recall

$$v_s(x) = \Psi\big(sF(x)\big), \qquad \sigma_s(x,\boldsymbol{\omega}) = \big|\boldsymbol{\omega}\cdot\nabla \log v_s(x)\big|,$$

where F is given by the HRBF model:

$$F(x) = \sum_{j=1}^{N} \Big(\alpha_j\, \varphi_j(x) - \boldsymbol{\beta}_j^{\top} \nabla\varphi_j(x)\Big) + p_0, \qquad (1)$$

and Ψ is a C^1 strictly increasing CDF with $\Psi(-\infty) = 0$ and $\Psi(+\infty) = 1$.

For a ray $\mathbf{r}(t) = \mathbf{o} + t\,\boldsymbol{\omega}$, $t \in [t_n, t_f]$, we set

$$C_s(\mathbf{r}) = \int_{t_n}^{t_f} c\big(\mathbf{r}(t),\boldsymbol{\omega}\big)\, \sigma_s\big(\mathbf{r}(t)\big)\, T_s(t)\, dt, \qquad (2)$$

$$T_s(t) = \exp\Big(-\int_{t_n}^{t} \sigma_s\big(\mathbf{r}(u)\big)\, du\Big), \qquad (3)$$

and assume $0 \le c(\mathbf{r}(t),\boldsymbol{\omega}) \le L$. For brevity, we write

$$f(t) = F(\mathbf{r}(t)), \qquad v_s(t) = v_s\big(\mathbf{r}(t)\big),$$
$$\sigma_s(t) = \sigma_s\big(\mathbf{r}(t),\boldsymbol{\omega}\big), \qquad c(t) = c\big(\mathbf{r}(t),\boldsymbol{\omega}\big).$$

We first prove two lemmas controlling the contribution along a ray, then combine them to obtain Theorem 1.

Lemma 1. *Let $[a,b] \subset [t_n, t_f]$ be an interval on which v_s is monotone, and assume $0 \le c(t) \le L$ on $[a,b]$. Define*

$$W_s(a,b) = T_s(a)\Big(1 - \frac{\min\{v_s(a), v_s(b)\}}{\max\{v_s(a), v_s(b)\}}\Big).$$

Then

$$\int_a^b c(t)\,\sigma_s(t)\,T_s(t)\,dt \;\le\; L\,W_s(a,b).$$

Proof. Along the ray,

$$\frac{d}{dt}\log v_s\big(\mathbf{r}(t)\big) = \nabla \log v_s\big(\mathbf{r}(t)\big)\cdot\boldsymbol{\omega},$$

so $\sigma_s(t) = \big|\frac{d}{dt}\log v_s(\mathbf{r}(t))\big|$, and $T_s'(t) = -\sigma_s(t)\,T_s(t)$.

Since v_s is monotone on $[a,b]$, $\log v_s$ is monotone as well, so there exists $\varepsilon \in \{\pm 1\}$ such that

$$\frac{d}{dt}\log v_s\big(\mathbf{r}(t)\big) = \varepsilon\,\sigma_s(t) \quad \text{on } [a,b].$$

Increasing case. If v_s is increasing, then $\varepsilon = +1$ and $v_s'(t) = v_s(t)\sigma_s(t)$, so

$$\frac{d}{dt}\big(v_s(t)T_s(t)\big) = 0, \quad \Rightarrow \quad v_s(t)T_s(t) = v_s(a)T_s(a),$$

and $T_s(t) = v_s(a)T_s(a)/v_s(t)$. With $u(t) = \log v_s(t)$ and $du = \sigma_s(t)\,dt$,

$$\int_a^b \sigma_s(t)\,T_s(t)\,dt = \int_{u(a)}^{u(b)} v_s(a)T_s(a)e^{-u}\,du$$
$$= T_s(a)\Big(1 - \frac{v_s(a)}{v_s(b)}\Big).$$

Here $\min\{v_s(a), v_s(b)\} = v_s(a)$ and $\max\{v_s(a), v_s(b)\} = v_s(b)$, so the right-hand side equals $W_s(a,b)$.

Decreasing case. If v_s is decreasing, then $\varepsilon = -1$ and $v_s'(t) = -v_s(t)\sigma_s(t)$. A direct computation gives

$$\frac{d}{dt}\left(\frac{T_s(t)}{v_s(t)}\right) = 0, \quad \Rightarrow \quad T_s(t) = T_s(a)\,\frac{v_s(t)}{v_s(a)}.$$

Using $\sigma_s(t)\,dt = -dv_s(t)/v_s(t)$,

$$\int_a^b \sigma_s(t)\,T_s(t)\,dt = -\frac{T_s(a)}{v_s(a)}\int_a^b dv_s(t)$$
$$= T_s(a)\Big(1 - \frac{v_s(b)}{v_s(a)}\Big),$$

and now $\min\{v_s(a), v_s(b)\} = v_s(b)$, $\max\{v_s(a), v_s(b)\} = v_s(a)$, so again $\int_a^b \sigma_s T_s = W_s(a, b)$.

Finally,

$$\int_a^b c(t)\,\sigma_s(t)\,T_s(t)\,dt \le L \int_a^b \sigma_s(t)\,T_s(t)\,dt = L\,W_s(a, b).$$

□

We now restrict the integral to where the ray intersects the HRBF support. By definition of the C^2 Wendland functions, each basis φ_j is radial and compactly supported, so $\operatorname{supp}\varphi_j$ is a closed ball $\overline{B(\boldsymbol{\mu}_j, R_j)}$. Let $S = \bigcup_{j=1}^N \operatorname{supp}\varphi_j$ and

$$I = \{t \in [t_n, t_f] : \mathbf{r}(t) \in S\}.$$

Lemma 2. *There exist $K \ge 0$ and points*

$$t_n \le a_1 < b_1 < \cdots < a_K < b_K \le t_f$$

such that $[a_k, b_k] \subset I$, the intervals $[a_k, b_k]$ are pairwise disjoint, and v_s is monotone on each $[a_k, b_k]$. Moreover,

$$C_s(\mathbf{r}) = \sum_{k=1}^K \int_{a_k}^{b_k} c(t)\,\sigma_s(t)\,T_s(t)\,dt.$$

Proof. If $\mathbf{r}(t) \notin S$, then $\varphi_j(\mathbf{r}(t)) = \nabla\varphi_j(\mathbf{r}(t)) = 0$ for all j, hence $F(\mathbf{r}(t)) = p_0$, $\nabla F(\mathbf{r}(t)) = 0$, so v_s is constant and $\sigma_s(t) = 0$. Therefore

$$C_s(\mathbf{r}) = \int_I c(t)\,\sigma_s(t)\,T_s(t)\,dt.$$

Each $\operatorname{supp}\varphi_j$ is a ball, so its intersection with the ray is either empty or a closed interval in $[t_n, t_f]$. The union of finitely many such intervals is a finite union of closed intervals, so

$$I = \bigcup_{\ell=1}^M [\alpha_\ell, \beta_\ell], \quad \alpha_\ell < \beta_\ell,$$

with these intervals chosen pairwise disjoint (merging overlaps if needed). Then

$$C_s(\mathbf{r}) = \sum_{\ell=1}^M \int_{\alpha_\ell}^{\beta_\ell} c(t)\,\sigma_s(t)\,T_s(t)\,dt.$$

Fix one such interval $[\alpha_\ell, \beta_\ell]$. The function $f(t) = F(\mathbf{r}(t))$ is a finite linear combination of $\varphi_j(\mathbf{r}(t))$ and $\nabla\varphi_j(\mathbf{r}(t))$, which are polynomial in t on this chord. Hence f and f' are polynomials on $[\alpha_\ell, \beta_\ell]$, and f' has only finitely many zeros unless it is identically zero. Since

$$v_s'(t) = \Psi'\big(s f(t)\big)\, s\, f'(t) \quad \text{with} \quad \Psi'(y) > 0,$$

the zeros of v_s' on $[\alpha_\ell, \beta_\ell]$ are exactly the zeros of f'. Thus, there exist finitely many points

$$\alpha_\ell = t_{\ell,0} < t_{\ell,1} < \cdots < t_{\ell,m_\ell} = \beta_\ell$$

such that on each $(t_{\ell,j-1}, t_{\ell,j})$, either v_s' is identically zero (so v_s is constant and $\sigma_s = 0$ there), or v_s' keeps a constant sign and v_s is strictly monotone.

Discarding the constant segments (which do not contribute), we obtain for this ℓ a finite family of subintervals

$$[a_{\ell,1}, b_{\ell,1}], \ldots, [a_{\ell,K_\ell}, b_{\ell,K_\ell}] \subset [\alpha_\ell, \beta_\ell]$$

on which v_s is monotone and

$$\int_{\alpha_\ell}^{\beta_\ell} c(t)\,\sigma_s(t)\,T_s(t)\,dt = \sum_{j=1}^{K_\ell} \int_{a_{\ell,j}}^{b_{\ell,j}} c(t)\,\sigma_s(t)\,T_s(t)\,dt.$$

Repeating this for all ℓ and renumbering all these monotone intervals as $[a_1, b_1], \ldots, [a_K, b_K]$, we obtain pairwise disjoint $[a_k, b_k] \subset I$ on which v_s is monotone and such that the integrand vanishes outside $\bigcup_{k=1}^K [a_k, b_k]$. This gives

$$C_s(\mathbf{r}) = \sum_{k=1}^K \int_{a_k}^{b_k} c(t)\,\sigma_s(t)\,T_s(t)\,dt.$$

□

We now prove Theorem 1 by combining the previous lemmas.

Theorem 1. *There exist finitely many pairwise disjoint intervals $[a_1, b_1], \ldots, [a_K, b_K] \subset [t_n, t_f]$, contained in the union of the Wendland supports and on which v is monotone, such that*

$$C(r) = \sum_{k=1}^K \int_{a_k}^{b_k} c(t)\,\sigma(t)\,T(t)\,dt.$$

If, for some k, $F(r(t)) > 0$ for all $t \in [a_k, b_k]$, then, as $s \to \infty$,

$$\int_{a_k}^{b_k} c(t)\,\sigma(t)\,T(t)\,dt \to 0.$$

Let

$$J = \{\, k : \exists\, t \in [a_k, b_k] \text{ with } F(r(t)) \le 0 \,\}.$$

Then

$$\lim_{s\to\infty} C(r) = \sum_{k\in J} \lim_{s\to\infty} \int_{a_k}^{b_k} c(t)\,\sigma(t)\,T(t)\,dt,$$

so the limiting color is a finite sum of contributions over intervals where v is monotone and along which the ray intersects the implicit surface or passes through the interior of the volume.

Proof. The decomposition

$$C_s(r) = \sum_{k=1}^{K} \int_{a_k}^{b_k} c(t)\,\sigma_s(t)\,T_s(t)\,dt$$

is exactly Lemma 2 applied to the ray r.

Fix k such that $F(r(t)) > 0$ for all $t \in [a_k, b_k]$. Since $t \mapsto F(r(t))$ is continuous and strictly positive on the compact interval $[a_k, b_k]$, there exists $\delta_k > 0$ with

$$F(r(t)) \geq \delta_k \quad \text{for all } t \in [a_k, b_k].$$

Set $f(t) = F(r(t))$ and $v_s(t) = \Psi(sf(t))$. Because $f(t) \geq \delta_k > 0$ and $\Psi(+\infty) = 1$, we have

$$v_s(t) = \Psi\big(sf(t)\big) \to 1 \quad \text{for all } t \in [a_k, b_k] \text{ as } s \to \infty.$$

In particular,

$$v_s(a_k) \to 1, \qquad v_s(b_k) \to 1,$$

so

$$\frac{\min\{v_s(a_k), v_s(b_k)\}}{\max\{v_s(a_k), v_s(b_k)\}} \longrightarrow 1.$$

By Lemma 1, on $[a_k, b_k]$ we have

$$\int_{a_k}^{b_k} c(t)\,\sigma_s(t)\,T_s(t)\,dt \leq L\,W_s(a_k, b_k),$$

with

$$W_s(a_k, b_k) = T_s(a_k)\left(1 - \frac{\min\{v_s(a_k), v_s(b_k)\}}{\max\{v_s(a_k), v_s(b_k)\}}\right).$$

We have $0 < T_s(a_k) \leq 1$, and the ratio in parentheses tends to 0, so $W_s(a_k, b_k) \to 0$ as $s \to \infty$, and therefore

$$\int_{a_k}^{b_k} c(t)\,\sigma_s(t)\,T_s(t)\,dt \to 0.$$

Thus, for every k such that $F(r(t)) > 0$ on $[a_k, b_k]$, the corresponding integral tends to 0 as $s \to \infty$.

Now $J = \{k : \exists t \in [a_k, b_k] \text{ with } F(r(t)) \leq 0\}$. For $k \notin J$, we have $F(r(t)) > 0$ on $[a_k, b_k]$, so the associated integral vanishes in the limit. Starting from

$$C_s(r) = \sum_{k=1}^{K} \int_{a_k}^{b_k} c(t)\,\sigma_s(t)\,T_s(t)\,dt,$$

we split $C_s(r)$:

$$\sum_{k\in J} \int_{a_k}^{b_k} c(t)\,\sigma_s(t)\,T_s(t)\,dt + \sum_{k\notin J} \int_{a_k}^{b_k} c(t)\,\sigma_s(t)\,T_s(t)\,dt.$$

The second sum tends to 0 as $s \to \infty$, since it has finitely many terms, each converging to 0. Hence

$$\lim_{s\to\infty} C_s(r) = \sum_{k\in J} \lim_{s\to\infty} \int_{a_k}^{b_k} c(t)\,\sigma_s(t)\,T_s(t)\,dt,$$

which is the claimed result. □

Theorem 1 is stated and proved under the assumption that $F(r(t)) > 0$ on $[a_k, b_k]$. The result shows that only intervals for which $F(r(t)) \leq 0$ can contribute in the limit $s \to \infty$.

In practice, this implies that if $p_0 \leq 0$ and $F \leq 0$ on the boundaries of Wendland supports, artificial interfaces may become visible. This observation motivates the choice $p_0 > 0$ adopted in the main formulation.

2. Implicit Function Initialization: Closed-form HRBF initialization

In the block-diagonal approximation, each 4×4 block associated with sample j is

$$D_{jj} = \begin{pmatrix} \varphi_j(\boldsymbol{\mu}_j) & -\nabla\varphi_j(\boldsymbol{\mu}_j)^\top \\ \nabla\varphi_j(\boldsymbol{\mu}_j) & -H\varphi_j(\boldsymbol{\mu}_j) \end{pmatrix}, \tag{4}$$

where $\varphi_j(x) = \varphi\big(\|x - \boldsymbol{\mu}_j\|/R_j\big)$. For the Wendland and Gaussian kernels used in this work we have $\varphi(0) = 1$ and $\nabla\varphi_j(\boldsymbol{\mu}_j) = \mathbf{0}$, while a second-order expansion around $\boldsymbol{\mu}_j$ yields an isotropic Hessian

$$H\varphi_j(\boldsymbol{\mu}_j) = \frac{\varphi''(0)}{R_j^2}\, I_3. \tag{5}$$

Adding Tikhonov regularization leads to

$$D_{jj} + \eta I_4 = \begin{pmatrix} 1+\eta & \mathbf{0}^\top \\ \mathbf{0} & \big(\eta - \frac{\varphi''(0)}{R_j^2}\big) I_3 \end{pmatrix}. \tag{6}$$

Solving $(D_{jj} + \eta I_4)\,\tilde{\boldsymbol{\lambda}}_j = \boldsymbol{b}_j$ with $\boldsymbol{b}_j = [-p_0,\, \boldsymbol{n}_j^\top]^\top$ gives

$$\tilde{\alpha}_j = -\frac{p_0}{1+\eta}, \qquad \tilde{\boldsymbol{\beta}}_j = \frac{1}{\eta - \frac{\varphi''(0)}{R_j^2}}\,\boldsymbol{n}_j. \tag{7}$$

For the Wendland kernel $\varphi_{\mathrm{W}}(r) = (1-r)_+^4(4r+1)$ we have $\varphi''_{\mathrm{W}}(0) = -20$, hence

$$\tilde{\alpha}_j^{\mathrm{W}} = -\frac{p_0}{1+\eta}, \qquad \tilde{\boldsymbol{\beta}}_j^{\mathrm{W}} = \frac{1}{\eta + 20/R_j^2}\,\boldsymbol{n}_j. \tag{8}$$

For the Gaussian kernel $\varphi_{\mathrm{G}}(r) = \exp(-\frac{1}{2}r^2)$ we have $\varphi''_{\mathrm{G}}(0) = -1$, yielding

$$\tilde{\alpha}_j^{\mathrm{G}} = -\frac{p_0}{1+\eta}, \qquad \tilde{\boldsymbol{\beta}}_j^{\mathrm{G}} = \frac{1}{\eta + 1/R_j^2}\,\boldsymbol{n}_j. \tag{9}$$

Here $\eta \geq 0$ is the regularization parameter: larger values of η increase the diagonal terms in $D_{jj} + \eta I_4$, shrinking $\tilde{\alpha}_j$ and $\tilde{\boldsymbol{\beta}}_j$ towards zero and thus controlling the amount of regularization in the HRBF formulation. In practice, we fix η based on the expected noise level of the dataset being evaluated.

DTU	Fast Dipole Sums	PGSR	Ours
24	28.13	31.36	**31.47**
37	26.03	27.56	**29.89**
40	28.01	30.60	**32.08**
55	29.47	**34.24**	32.27
63	29.68	32.13	**32.44**
65	**31.83**	31.63	28.89
69	29.35	**31.54**	26.51
83	30.15	**34.61**	32.24
97	28.81	**31.35**	29.06
105	30.80	**32.68**	31.48
106	34.42	**36.92**	26.48
110	22.10	**35.37**	28.55
114	31.23	**32.44**	31.24
118	35.71	**39.24**	29.05
122	36.00	**39.57**	34.36
Mean	30.11	**33.42**	30.40

Table 1. PSNR (dB) on the DTU [1] dataset for Fast Dipole Sums [3], PGSR [2], and our method. Higher is better.

Scene	**Fast Dipole Sums**	**Ours**
basketball	19.83	**22.18**
bear	25.03	**26.22**
bread	29.25	**30.18**
camera	21.95	**22.48**
clock	25.92	**28.38**
cow	23.86	**26.92**
dog	27.22	**28.89**
doll	**23.42**	23.39
dragon	**27.86**	26.98
durian	24.12	**24.69**
fountain	21.26	**29.98**
gundam	23.72	**26.86**
house	27.75	**29.43**
jade	29.66	**30.85**
man	**34.86**	32.69
monster	26.00	**27.05**
sculpture	26.29	**32.31**
stone	23.46	**26.65**
Mean	25.64	**27.56**

Table 2. PSNR (dB) on the BlendedMVS [11] dataset for Fast Dipole Sums [3] and our method. Higher is better.

3. Analysis of the Radiance Field Representation

The goal of this section is to analyze our radiance parameterization. We first report quantitative rendering results on the DTU and BlendedMVS datasets by evaluating the PSNR of the generated images. We then study the influence of using spherical harmonics and spherical Gaussians for the radiance representation, and finally analyze the relationship between our radiance field and the implicit function defining the geometry.

3.1. Quantitative Rendering Evaluation

We first evaluate the rendering quality of our approach on DTU [1] and BlendedMVS [11], measuring the peak signal-to-noise ratio (PSNR) between rendered images and ground-truth views. Table 1 reports PSNR values on DTU for Fast Dipole Sums [3], PGSR [2], and our method. While PGSR achieves the highest mean PSNR (33.42 dB), our approach obtains 30.40 dB, slightly above Fast Dipole Sums (30.11 dB). Despite this numerical gap, our algorithm produces photorealistic renderings of the scenes: the qualitative results in Figure 2 (and the corresponding meshes in Figure 1) show that our outputs are visually comparable to PGSR. This discrepancy between PSNR and perceived quality may be partly explained by illumination changes across DTU views and by the significant noise present in the foreground masks, which both tend to penalize pixel-wise metrics.

On BlendedMVS, we compare our method to Fast Dipole Sums in Table 2. Our approach consistently improves the PSNR across most small-object scenes, yielding a higher mean PSNR (27.56 dB vs. 25.64 dB) on this dataset.

3.2. Radiance Parameterization with Spherical Harmonics and Spherical Gaussians

We now study the impact of the radiance parameterization on both rendering quality and geometric accuracy on the BlendedMVS dataset [11]. Table 3 compares three variants: using only spherical harmonics (SH), only spherical Gaussians (SG), or a hybrid parameterization combining SH and SG. We report PSNR on rendered images and Chamfer Distance (CD) on the reconstructed geometry.

Using SH only yields the highest PSNR (27.92 dB) but a relatively larger CD (0.58), indicating slightly less accurate geometry. In contrast, using SG only produces the best CD (0.51), but at the cost of a lower PSNR (27.01 dB). The hybrid SH–SG parameterization achieves a balanced trade-off, with PSNR (27.56 dB) close to the SH-only variant and CD (0.52) close to the SG-only variant. We therefore adopt this combined SH–SG radiance parameterization as our final choice, as it offers a good compromise between rendering quality and reconstruction accuracy on BlendedMVS.

3.3. Coupling Between Radiance Field and Geometry

In this subsection, we analyze the relation of the radiance field to the underlying HRBF-based implicit geometry, and

Radiance representation	PSNR ↑ (dB)	CD ↓
SH only	**27.92**	0.58
SG only	27.01	**0.51**
SH + SG	27.56	0.52

Table 3. Ablation on radiance representations (spherical harmonics, spherical Gaussians, and their combination), reported in terms of PSNR and Chamfer Distance (CD).

study an alternative formulation in which radiance more directly reflects the influence of individual primitives on the implicit field.

In Section 4.3, *Radiance Field Representation*, we model the radiance field c as

$$c(\boldsymbol{\omega}, \mathbf{x}) = \sum_{m=1}^{N} w_m(\mathbf{x})\, c_m(\boldsymbol{\omega}), \tag{10}$$

where $\boldsymbol{\omega} \in \mathbb{S}^2$ is the viewing direction, $\mathbf{x} \in \mathbb{R}^3$ is the query position, and

$$w_m(\mathbf{x}) = \frac{\varphi(\|\mathbf{x} - \boldsymbol{\mu}_m\|/R_m)}{\sum_{j=1}^{N} \varphi(\|\mathbf{x} - \boldsymbol{\mu}_j\|/R_j)}, \tag{11}$$

with φ a radial basis function centered at $\boldsymbol{\mu}_m$ and scale R_m. This representation yields a simple, differentiable formulation in which the radiance at any point is expressed as a weighted average of local primitive radiance fields, with the influence of each primitive controlled by its distance to the query point.

The volume density $\sigma(\mathbf{x})$ is directly derived from our surface representation, which we model using an HRBF-based implicit function:

$$F(\mathbf{x}) = \sum_{j=1}^{N} \Big(\alpha_j\, \varphi_j(\mathbf{x}) - \boldsymbol{\beta}_j^{\top} \nabla \varphi_j(\mathbf{x}) \Big) + p_0, \tag{12}$$

where each basis function φ_j is an RBF with center $\boldsymbol{\mu}_j$ and scale R_j. While the centers $\boldsymbol{\mu}_j$ and scales R_j control the location and spatial extent of the radial basis functions, the behavior of F within their support is mainly governed by the HRBF coefficients α_j and $\boldsymbol{\beta}_j$. For a given primitive j, its contribution to the implicit function is

$$F_j(\mathbf{x}) = \alpha_j\, \varphi_j(\mathbf{x}) - \boldsymbol{\beta}_j^{\top} \nabla \varphi_j(\mathbf{x})\,. \tag{13}$$

In Eq. (10), the radiance c at a given point does not depend on α_j or $\boldsymbol{\beta}_j$. Consequently, the parameters that control the implicit function, and thus the density, do not directly influence the radiance. We refer to this design as a *decoupled formulation*.

This decoupling offers two practical advantages: it extrapolates radiance to any point in space using the same RBF kernels as the implicit function, and it remains easily differentiable, which is convenient for optimization. It also implies that the local influence of a primitive on the variations of the implicit function is not correlated with its influence on the radiance (except through the distance-dependent RBF weight). Hence, a primitive that induces strong variations of F and therefore generates high density locally (since the density depends on the variations of the implicit function near the isosurface $F(\mathbf{x}) = 0$) does not, for the same distance and scale, contribute more to the radiance than a primitive that only weakly affects the implicit function.

Coupled radiance. We now introduce an alternative *coupled* radiance formulation which makes the radiance explicitly depend on the HRBF weighting parameters. Our goal is that a primitive which induces large variations of the implicit field F should, for the same distance and scale, have a stronger influence on the radiance than a primitive that only weakly contributes to F. This encourages the model to associate the primitive responsible for the local surface geometry with the explanation of the local radiance as well.

We write the implicit field as

$$F(\mathbf{x}) = p_0 + \sum_{j} F_j(\mathbf{x}), \tag{14}$$

where F_j denotes the contribution of primitive j, and the implicit surface is defined as the zero level set $F(\mathbf{x}) = 0$. To couple radiance and geometry, we introduce a scalar criterion that quantifies the influence of primitive j on the field:

$$\mathcal{I}_j \;:=\; \|F_j\|^2_{L^2(\Omega_j)} \;=\; \int_{\Omega_j} F_j(\mathbf{x})^2\, d\mathbf{x}, \tag{15}$$

where Ω_j is a spatial domain associated with primitive j. Intuitively, $\mathcal{I}_j$ measures how strongly primitive j deviates $F(\mathbf{x})$ from the background value p_0 in its neighborhood and thus how much it can affect the shape of the zero level set.

Since each primitive is spatially localized, we use a normalized influence measure obtained by dividing by the volume of its effective support. Let Ω_j be a compact region that captures the support of primitive j (for Gaussian RBF, we use a ball of radius $3\sigma_j$ centered at $\boldsymbol{\mu}_j$), and let $V_j = \mathrm{Vol}(\Omega_j)$. We then define the average influence per unit volume as

$$\bar{\mathcal{I}}_j \;:=\; \frac{1}{V_j} \int_{\Omega_j} F_j(\mathbf{x})^2\, d\mathbf{x}. \tag{16}$$

This normalized criterion favors primitives that induce strong variations of F relative to the background value p_0, and the normalization by V_j makes the score comparable across primitives with different spatial supports.

In the coupled radiance formulation, we reuse this normalized influence as a multiplicative factor in the spatial weights of Eq. (11). Concretely, we obtain the influence-modulated weights

$$\tilde{w}_m(\mathbf{x}) = \frac{\bar{\mathcal{I}}_m \, \varphi(\|\mathbf{x} - \boldsymbol{\mu}_m\|/R_m)}{\sum_{j=1}^{N} \bar{\mathcal{I}}_j \, \varphi(\|\mathbf{x} - \boldsymbol{\mu}_j\|/R_j)}. \tag{17}$$

With this modification, primitives that induce stronger overall variations of the implicit field F (according to $\bar{\mathcal{I}}_m$) are given proportionally more weight in the radiance blending.

Gaussian Radial Basis Function. Since our best results in the main paper are obtained with the Gaussian kernel, we focus on this case and derive the closed-form expression of the normalized influence criterion $\bar{\mathcal{I}}_j$. Each primitive j is parameterized by a center $\boldsymbol{\mu}_j \in \mathbb{R}^3$, a scale $\sigma_j > 0$, a scalar weight $\alpha_j \in \mathbb{R}$, and a vector weight $\boldsymbol{\beta}_j \in \mathbb{R}^3$. The Gaussian kernel is defined as

$$G_j(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \boldsymbol{\mu}_j\|^2}{2\sigma_j^2}\right), \tag{18}$$

and the contribution of primitive j to the field is

$$F_j(\mathbf{x}) = \alpha_j \, G_j(\mathbf{x}) - \langle \boldsymbol{\beta}_j, \nabla G_j(\mathbf{x}) \rangle. \tag{19}$$

Using the closed-form gradient

$$\nabla G_j(\mathbf{x}) = -\frac{1}{\sigma_j^2} \, G_j(\mathbf{x}) \, (\mathbf{x} - \boldsymbol{\mu}_j), \tag{20}$$

we can rewrite

$$F_j(\mathbf{x}) = \left(\alpha_j + \frac{1}{\sigma_j^2} \langle \boldsymbol{\beta}_j, \mathbf{x} - \boldsymbol{\mu}_j \rangle\right) G_j(\mathbf{x}). \tag{21}$$

For Gaussian primitives, the integral of $F_j(\mathbf{x})^2$ over the whole space $\mathbb{R}^3$ admits a closed form:

$$\int_{\mathbb{R}^3} F_j(\mathbf{x})^2 \, d\mathbf{x} = \pi^{3/2} \left(\alpha_j^2 \, \sigma_j^3 \; + \; \frac{1}{2} \, \|\boldsymbol{\beta}_j\|^2 \, \sigma_j\right). \tag{22}$$

During the evaluation of the implicit function, we use a truncated Gaussian RBF with finite support. Consistently, we define the normalized influence over the truncation ball

$$\Omega_j = \{\mathbf{x} : \|\mathbf{x} - \boldsymbol{\mu}_j\| \le 3\sigma_j\},$$

which we use as the effective support of primitive j. Since $F_j(\mathbf{x})^2$ is a Gaussian up to a low-degree polynomial factor, its tails decay exponentially fast, and the contribution of the complement $\mathbb{R}^3 \setminus \Omega_j$ to $\int_{\mathbb{R}^3} F_j(\mathbf{x})^2 \, d\mathbf{x}$ is negligible. We therefore approximate

$$\int_{\Omega_j} F_j(\mathbf{x})^2 \, d\mathbf{x} \;\approx\; \int_{\mathbb{R}^3} F_j(\mathbf{x})^2 \, d\mathbf{x}. \tag{23}$$

Using the volume of the ball $V_j = \mathrm{Vol}(\Omega_j) = 36\pi\sigma_j^3$, the normalized influence per unit volume

$$\bar{\mathcal{I}}_j \;=\; \frac{1}{V_j} \int_{\Omega_j} F_j(\mathbf{x})^2 \, d\mathbf{x} \tag{24}$$

admits the closed-form approximation

$$\bar{\mathcal{I}}_j \;\approx\; \frac{\sqrt{\pi}}{36} \left(\alpha_j^2 \;+\; \frac{\|\boldsymbol{\beta}_j\|^2}{2\sigma_j^2}\right). \tag{25}$$

This provides a simple, comparable, and easily differentiable scalar score across primitives that we use to modulate the coupled radiance weights.

In practice, we did not observe significant differences between the decoupled and coupled formulations in terms of rendering quality or surface reconstruction on the BlendedMVS dataset [11]. For this reason, we retain the simpler decoupled radiance parameterization in our main experiments and present the coupled variant here primarily for its theoretical interest. A more in-depth investigation of radiance parameterizations that explicitly depend on the underlying implicit geometry is left for future work.

4. Details on evaluation protocol on DTU dataset [1]

The DTU dataset (used by recent surface reconstruction methods such as NeuS [9], 2DGS [5], Neuralangelo [6], Fast Dipole Sums [3], and PGSR [2]) comprises 15 scenes selected from the 80 scenes of the full DTU dataset (released in 2014). These 15 scenes were chosen by the authors of IDR [12] (2020), who also provided manually annotated foreground masks (three masks were previously annotated in [8]). Each scene contains either 49 or 64 images at 1600×1200 pixels. Camera poses are accurately determined using a robotic arm and a calibration target. Ground-truth geometry is captured with a structured-light scanner. Earlier methods such as NeuS [9] and Neuralangelo [6] relied on full-resolution RGB images (1600×1200 pixels). In contrast, 2DGS [5], one of the first Gaussian Splatting–based approaches, evaluated DTU at half-resolution (800×600 pixels). Subsequent Gaussian-Splatting-based methods (GOF [13], PGSR [2]) adopted the same 800×600 resolution. RaDe-GS [14] further showed that using lower-resolution inputs (800×600 instead of 1600×1200) can improve mesh-reconstruction performance for methods such as 2DGS and GOF. For consistency and fairness, we therefore use full-resolution images (1600×1200 pixels) and retrain Fast Dipole Sums [3] and PGSR [2] at this resolution. We selected these two approaches because they currently represent state-of-the-art surface reconstruction methods: Fast Dipole Sums uses an implicit dipole-based representation over points, while PGSR relies on an explicit representation derived from Gaussian Splatting. These recomputed

results correspond to the entries marked with * in Table 1 of the main paper. In these training procedures and during our method optimization, all images (at 1600×1200 resolution) are used, together with the foreground masks provided by IDR [12].

5. Details on evaluation protocol on BlendedMVS dataset [11]

BlendedMVS [11] is a large-scale MVS dataset comprising 113 scenes. The images are rendered from reconstructed meshes and blended with the original inputs to preserve view-dependent lighting. This dataset was first employed for surface reconstruction by NeuS [9], which selected 7 scenes. Gaussian Surfels [4] later defined a subset of 18 scenes that subsequently became the standard benchmark, adopted by Fast Dipole Sums [3] and by our work. Surface-reconstruction methods use the low-resolution images at 768×576 pixels. As in Gaussian Surfels [4], NeuS2 [10], Fast Dipole Sums [3], and our approach, all images are used during training. We retrained Fast Dipole Sums [3] on the BlendedMVS dataset. We chose this method because it currently represents the state of the art for this dataset. These results correspond to the entries marked with * in Table 2 of the main paper.

6. Additional qualitative results on DTU dataset [1] and BlendedMVS dataset [11]

Figures 1 and 2 provide additional mesh and rendering results on the DTU dataset [1]. Figures 3 and 4 show supplementary results on the BlendedMVS dataset [11]. In the main paper, mesh renderings were produced with Blender's Cycles engine, whose high-quality shading facilitates visual inspection of geometric details. However, because the DTU and BlendedMVS ground truth consists of point clouds, these cannot be rendered directly in Blender. In the supplementary material, we therefore rely on CloudCompare to compare reconstructed meshes with reference point clouds.

The figures indicate that Fast Dipole Sums [3] and PGSR [2] tend to generate meshes lacking some of the fine details captured by our method, albeit at the cost of slightly increased noise in our reconstructions. PGSR occasionally produces incorrect fillings in regions with poor visibility due to its reliance on depth-image integration (visible on the roof boundaries in DTU *scan_24*). Fast Dipole Sums, owing to its global formulation, may generate undesirable outlier surfaces (e.g., around the fruits in DTU *scan_63*). Regarding appearance rendering, there is a clear quality gap between Fast Dipole Sums and both PGSR and our method. The MLP-based renderer of Fast Dipole Sums produces blurrier color outputs, leading to a loss of detail (e.g., the eye of the stuffed toy in DTU *scan_105*), whereas our point-based formulation preserves texture fidelity.

Method	Barn	Caterpillar	Courthouse	Ignatius	Meetingroom	Truck	Mean	Time
Ours	0.69	0.39	0.26	0.81	0.31	0.56	0.503	1.3h
PGSR	0.66	0.41	0.21	0.80	0.29	0.60	0.495	1.2h

Table 4. Quantitative F1-score results on Tanks and Temples.

Finally, the reference point clouds in DTU used as ground truth contain holes (accounted for by observation masks during evaluation) as well as noticeable noise and outliers. This may ultimately impose an upper bound on the achievable evaluation accuracy on the DTU dataset.

7. Additional Quantitative Results on Tanks and Temples

Table 4 reports additional quantitative comparisons on the Tanks and Temples dataset using the F1-score metric. We compare our method against PGSR across six scenes: *Barn*, *Caterpillar*, *Courthouse*, *Ignatius*, *Meetingroom*, and *Truck*.

Overall, our method achieves a higher mean F1-score (0.503 vs. 0.495), demonstrating a consistent improvement over PGSR. Notably, we observe gains on scenes such as *Barn*, *Courthouse*, *Ignatius*, and *Meetingroom*. PGSR performs slightly better on *Caterpillar* and *Truck*, suggesting that certain geometric or texture characteristics of these scenes may favor its formulation. Nevertheless, our approach remains competitive and achieves superior overall performance.

In terms of computational cost, both methods exhibit comparable runtimes, with our method taking approximately 1.3 hours and PGSR 1.2 hours per scene. This highlights that the performance gain is obtained without significant additional computational overhead.

These results further support the robustness of our method across diverse real-world scenes.

References

[1] Henrik Aanæs, Rasmus RamsbØl Jensen, George Vogiatzis, Engin Tola, and Anders Bjorholm Dahl. Large-scale data for multiple-view stereopsis. *International Journal of Computer Vision*, 120(2):153–168, 2016. 1, 4, 6, 7, 8, 9

[2] Danpeng Chen, Hai Li, Weicai Ye, Yifan Wang, Weijian Xie, Shangjin Zhai, Nan Wang, Haomin Liu, Hujun Bao, and Guofeng Zhang. Pgsr: Planar-based gaussian splatting for efficient and high-fidelity surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–12, 2024. 4, 6, 7, 8, 9

[3] Hanyu Chen, Bailey Miller, and Ioannis Gkioulekas. 3d reconstruction with fast dipole sums. *ACM Transaction on Graphics (TOG)*, 43(6), 2024. 4, 6, 7, 8, 9, 10, 11

[4] Pinxuan Dai, Jiamin Xu, Wenxiang Xie, Xinguo Liu, Huamin Wang, and Weiwei Xu. High-quality surface reconstruction using gaussian surfels. In *Special Interest Group on Computer Graphics and Interactive Techniques (SIGGRAPH)*, 2024. 7

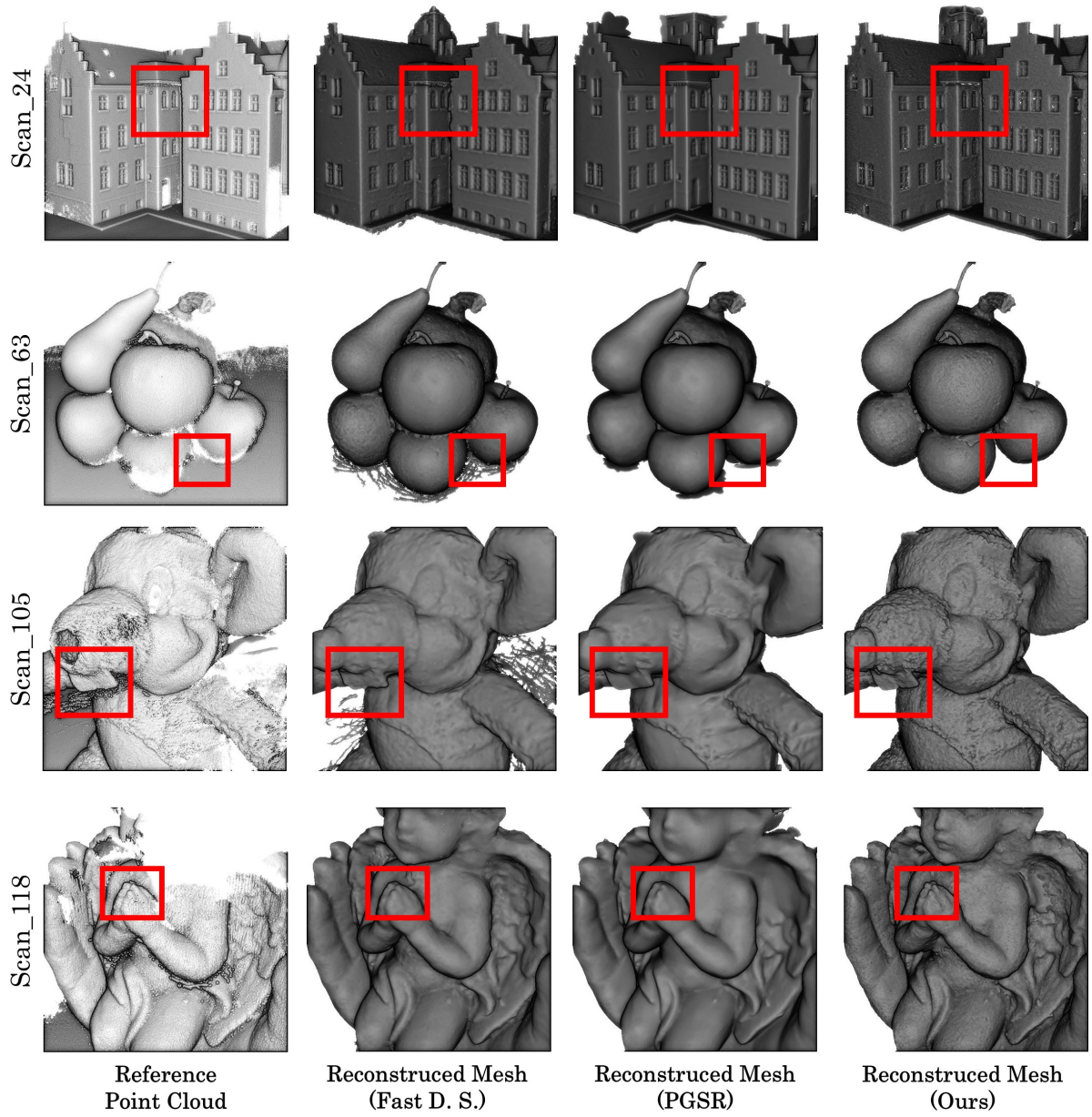

Figure 1. Qualitative meshing results on the DTU dataset [1] between Fast Dipole Sums [3], PGSR [2] and our approach.

[5] Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2d gaussian splatting for geometrically accurate radiance fields. In *Special Interest Group on Computer Graphics and Interactive Techniques (SIGGRAPH)*, 2024. 6

[6] Zhaoshuo Li, Thomas Müller, Alex Evans, Russell H. Taylor, Mathias Unberath, Ming-Yu Liu, and Chen-Hsuan Lin. Neuralangelo: High-fidelity neural surface reconstruction. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8456–8465, 2023. 6

[7] Bailey Miller, Hanyu Chen, Alice Lai, and Ioannis Gkioulekas. Objects as volumes: A stochastic geometry view of opaque solids. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 87–97, 2024. 1

[8] Michael Niemeyer, Lars Mescheder, Michael Oechsle, and Andreas Geiger. Differentiable volumetric rendering: Learning implicit 3d representations without 3d supervision. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3501–3512, 2020. 6

[9] Peng Wang, Lingjie Liu, Yuan Liu, Christian Theobalt, Taku Komura, and Wenping Wang. Neus: learning neural implicit surfaces by volume rendering for multi-view reconstruction.

Figure 2. Qualitative rendering results on the DTU dataset [1] between Fast Dipole Sums [3], PGSR [2] and our approach.

In *International Conference on Neural Information Processing Systems (NeurIPS)*, Red Hook, NY, USA, 2021. Curran Associates Inc. 6, 7

[10] Yiming Wang, Qin Han, Marc Habermann, Kostas Daniilidis, Christian Theobalt, and Lingjie Liu. Neus2: Fast learning of neural implicit surfaces for multi-view reconstruction. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023. 7

[11] Yao Yao, Zixin Luo, Shiwei Li, Jingyang Zhang, Yufan Ren, Lei Zhou, Tian Fang, and Long Quan. Blendedmvs: A large-scale dataset for generalized multi-view stereo networks. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1787–1796, 2020. 1, 4, 6, 7, 10, 11

[12] Lior Yariv, Yoni Kasten, Dror Moran, Meirav Galun, Matan Atzmon, Ronen Basri, and Yaron Lipman. Multiview neural surface reconstruction by disentangling geometry and appearance. In *International Conference on Neural Information Processing Systems (NeurIPS)*, Red Hook, NY, USA, 2020. 6, 7

[13] Zehao Yu, Torsten Sattler, and Andreas Geiger. Gaussian opacity fields: Efficient adaptive surface reconstruction in

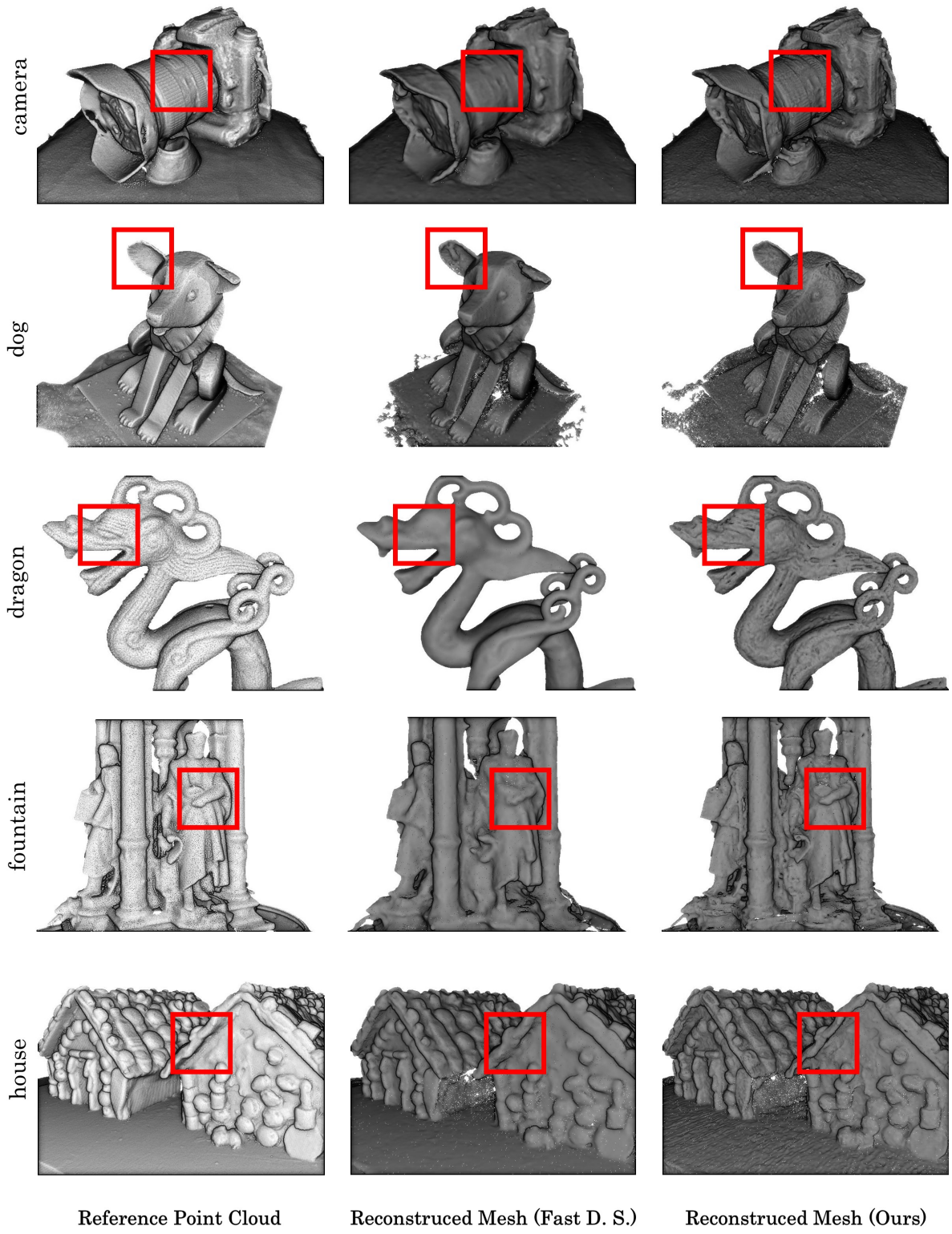

Figure 3. Qualitative meshing results on the BlendedMVS dataset [11] between Fast Dipole Sums [3] and our approach.

Figure 4. Qualitative rendering results on the BlendedMVS dataset [11] between Fast Dipole Sums [3] and our approach.

unbounded scenes. *ACM Transactions on Graphics (TOG)*, 43(6), 2024. 6

[14] Baowen Zhang, Chuan Fang, Rakesh Shrestha, Yixun Liang, Xiaoxiao Long, and Ping Tan. RaDe-GS: Rasterizing Depth in Gaussian Splatting. *arXiv e-prints*, art. arXiv:2406.01467, 2024. 6