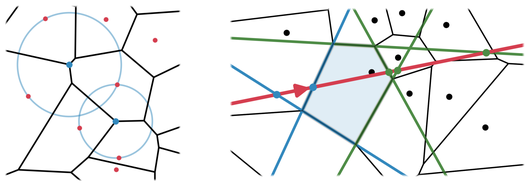
FoamSurf: High-Quality Surface Reconstruction Using RadFoam

Ignacio Dassori, Mats L. Grobe, Lukas Höllein

The Idea

RadFoam (Radiant Foam) is a differentiable 3D scene representation based on Voronoi tessellations.



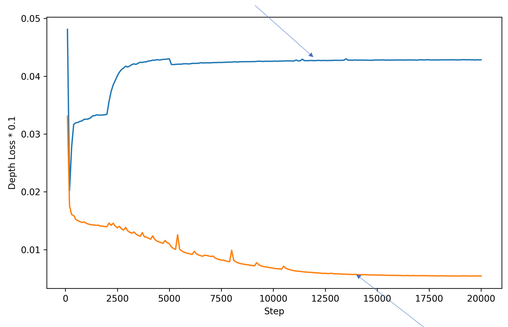
RadFoam's native quantile-based depth rendering was supplemented with **expected depth**. Face normal were retrieved via the Delaunay-Voronoi duality.

$$\hat{D} = \sum_{i \in N} d_i \, \alpha_i \prod_{j=1}^{i-1} (1 - \alpha_j) \qquad \widehat{N} = \sum_{i \in \mathcal{N}} \mathbf{n}_i \, \alpha_i \, T_i \qquad \mathbf{n}_i = \frac{p_{i+1} - p_i}{\|p_{i+1} - p_i\|}$$

CUDA Implementation

$$T_{j-1} = \prod_{i<j} (1 - \alpha_i), \quad t_j = \tfrac{1}{2}(t_{0,j} + t_{1,j}), \quad w_j = T_{j-1}\,\alpha_j, \quad D = \sum_{j=1}^{K} w_j \, t_j.$$

$$\frac{\partial D}{\partial \alpha_k} = \frac{\partial}{\partial \alpha_k}(T_{k-1}\,\alpha_k\,m_k) = T_{k-1}\,m_k$$



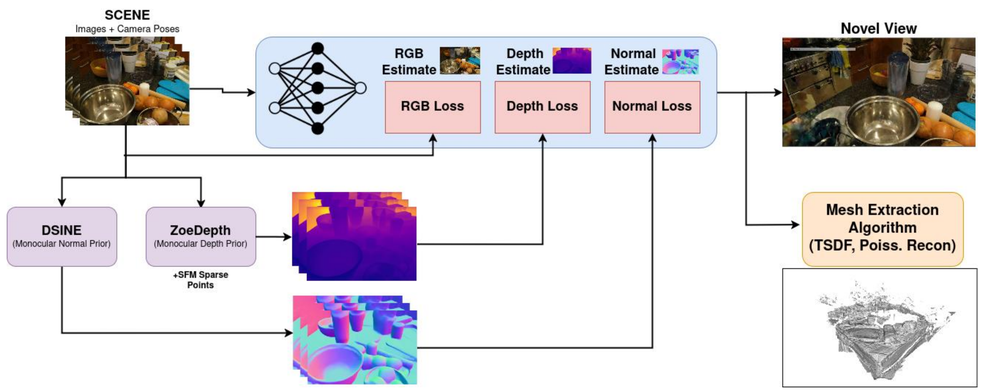
$$\boxed{\frac{\partial D}{\partial \alpha_k} = T_{k-1}\,m_k \; + \; \left(-\frac{S}{1-\alpha_k}\right) = T_{k-1}\left(m_k - \frac{S}{T_{k-1}(1-\alpha_k)}\right)}$$

$$S = \sum_{j>k} T_{j-1}\,\alpha_j\,m_j,$$

Code available at:

https://github.com/matsgrobe/FoamSurf

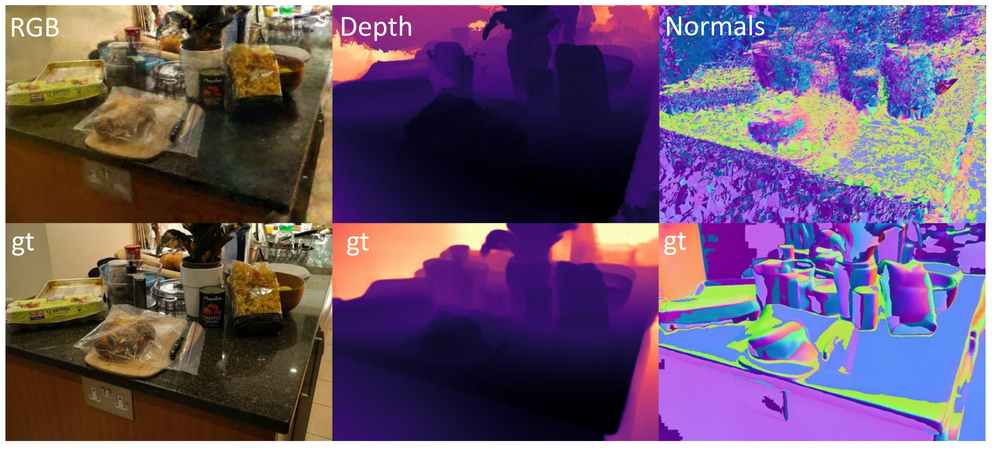
The Pipeline



Monocular Supervision

$$\mathcal{L}_{\hat{\mathbf{D}}} = \mathbf{g}_{\text{rgb}} \frac{1}{|\hat{\mathbf{D}}|} \sum \log\left(\mathbf{1} + ||\hat{\mathbf{D}} - \mathbf{D}||_1\right) \qquad \mathcal{L}_{\hat{\mathbf{N}}} = \frac{\mathbf{1}}{|\hat{\mathbf{N}}|} \sum \left(\mathbf{1} - \frac{\hat{\mathbf{N}} \cdot \mathbf{N}}{\|\hat{\mathbf{N}}\|_2 \|\mathbf{N}\|_2}\right) \qquad \mathcal{L} = \mathcal{L}_{\hat{\mathbf{C}}} + \lambda_{\mathbf{d}} \mathcal{L}_{\hat{\mathbf{D}}} + \lambda_{\mathbf{n}} \mathcal{L}_{\hat{\mathbf{N}}}$$

Edge Aware Depth Loss ***Cosine Loss*** ***Final Loss***



Scenes above trained on MipNerf360 with monocular priors and right column trained on NerfSynthetic with gt depths.

Table 1: Ablation (MipNerf360)

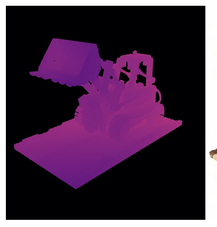
Loss variant	PSNR	SSIM	$RMSE_d$	$CosSim_n$
$L_{\hat{C}}$	**28.74**	**.8807**	1.272	52.07
$L_{\hat{C}} + L_{\hat{D}}$	21.91	.7087	**.9628**	50.47
$L_{\hat{C}} + L_{\hat{N}}$	21.30	.6096	27.391	**29.35**
$L_{\hat{C}} + L_{\hat{D}} + L_{\hat{N}}$	21.87	.6436	1.195	36.21
$L_{\hat{C}} + L_{\hat{D}} \rightarrow logL_1$	23.86	.7882	1.007	49.83

RadFoam

2DGS (downsampled)

2DGS

3D meshes are extracted from rendered RGB and depth maps using the **TSDF** algorithm.



FoamSurf

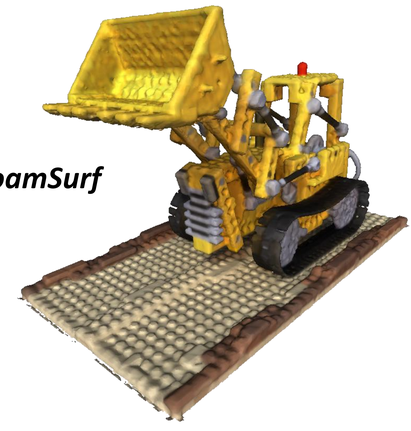


Table 2: Depth estimation and mesh evaluation (Synthetic)

Method	PSNR	RMSE	MAE	Chamfer
RadFoam	28.50	.1631	.0338	.0182
FoamSurf ($\log L_1$)	27.44	.1395	.0198	.0157
FoamSurf ($L_{\hat{D}}$)	24.34	.1088	**.0161**	**.0141**
2DGS	**37.75**	**.0903**	.0674	.2809
2DGS (downsampled)	29.32	.0984	.0728	.0221