

Supplementary Material of Free-SurGS: SfM-free 3D Gaussian Splatting for Surgical Scene Reconstruction

Anonymous

Anonymous Organization

1 Implementation Details and Results

Epipolar Geometry. In dynamic surgical environments characterized by transient objects and photometric inconsistencies, it is essential to identify and preserve correspondences that are both rigid and reliable for accurate matching. We compute the Sampson distance to measure the geometric error between a point in one image and its corresponding epipolar line in the other image:

$$d(\mathbf{x}, \mathbf{x}', \mathbf{F}) = \frac{(\mathbf{x}^T \mathbf{F} \mathbf{x})^2}{(\mathbf{F} \mathbf{x})_1^2 + (\mathbf{F} \mathbf{x})_2^2 + (\mathbf{F}^T \mathbf{x}')_1^2 + (\mathbf{F}^T \mathbf{x}')_2^2}, \quad (1)$$

where \mathbf{F} denotes the fundamental matrix of the estimated relative camera poses, \mathbf{x} and \mathbf{x}' are the homogeneous coordinates of matched points in the image pair, $(\mathbf{F} \mathbf{x})_1$ and $(\mathbf{F} \mathbf{x})_2$ denote the first and second component of the vector $\mathbf{F} \mathbf{x}$. By computing the Sampson distance, we take a threshold β to obtain a rigid mask $\mathbf{M}_r = d(\mathbf{x}, \mathbf{x}', \mathbf{F}) < \beta$ for time t , ensuring that only robust correspondences are utilized for subsequent pose estimation tasks from t to $t + 1$.

Scale-invariant Depth Loss. To improve the geometry of 3D Gaussian, we adopt the scale-invariant depth loss to use the monocular depth \mathbf{D} to supervise the rendered depth $\hat{\mathbf{D}}$. We first align the \mathbf{D} to $\hat{\mathbf{D}}$ to get $\tilde{\mathbf{D}}$:

$$\mathcal{L}_{dep} = \mathcal{L}_{reg} + \mathcal{L}_2 = \frac{1}{M} \|\tilde{\mathbf{D}} - \mathbf{D}\|_2^2 + \frac{1}{M} \sum_{k=1}^K \sum_{i=1}^M (|\nabla_x \mathbf{R}_i^k + \nabla_y \mathbf{R}_i^k|), \quad (2)$$

where \mathbf{R}_i denotes the difference between $\tilde{\mathbf{D}}$ and $\hat{\mathbf{D}}$ with scale level $K = 4$, M denotes the total pixels of image.

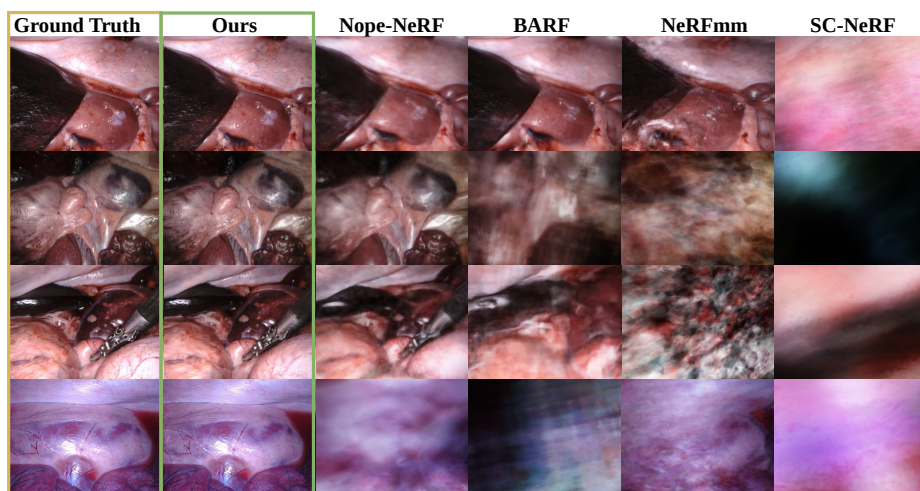


Fig. 1. Qualitative results of novel view synthesis on the SCARED Dataset.