SurgicalGaussian: Deformable 3D Gaussians for High-Fidelity Surgical Scene Reconstruction - Supplementary Material

1 GIDM initialization

In order to effectively provide a good initialization of Gaussian primitives, we project image to get 3D point cloud. However, the holes caused by removing the surgical instrument from the first frame are large, and Gaussian points cannot be placed in these occluded areas, which affects the reconstruction of tissue in this area. Therefore, we check what is visible in the subsequent frames to fill in the missing regions that are not visible in frame $\mathbf{I}_0$. As show in Fig. 1, we checked all frames and obtained the refined image $\mathbf{I}^*$ and depth map $\mathbf{D}^*$. The hole in $\mathbf{I}^*$ is the region that's invisible all the time. According to Eq.(2) in the main manuscript, we project $\mathbf{I}^*$ to get point cloud $\mathbf{P}^*$ and downsample it as the initialization of 3D Gaussians.

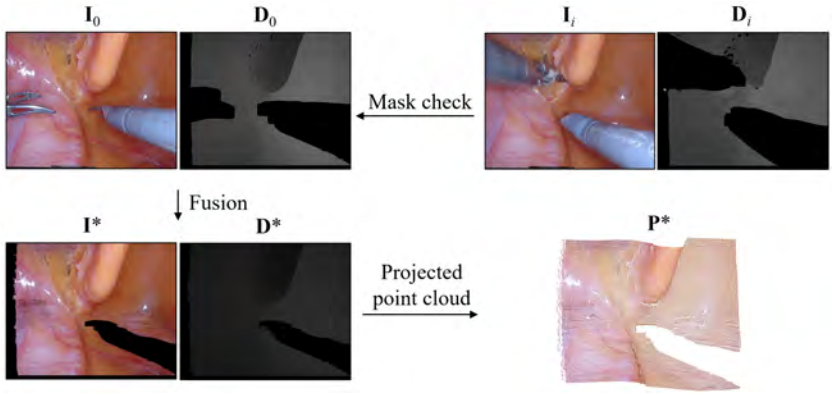

Fig. 1. Gaussian initialization strategy.

2 User study

In addition to quantitatively measuring the 3D reconstruction quality, we conducted a user survey to further validate the value of our method in real medical applications. We asked 10 surgical experts to evaluate different reconstruction methods by a questionnaire. In this study, we presented reconstruction videos and point clouds of different methods applied to four surgical scenes, accessible

on the project website. Initially, we concealed all identifying information about the videos and point clouds. Then, participants are asked to rate the video quality and point cloud quality on a scale from 1 to 5, with 1 being "Very poor" and 5 being "Excellent". The questionnaire we designed measures the reconstruction quality of surgical scenes from multiple aspects: recovery of texture details of organs, capture of soft tissue motion, and recovery of tissue occluded by surgical instruments. In addition, we provided reconstructed point clouds of the dynamic surgical scene and asked users to evaluate the point cloud quality.

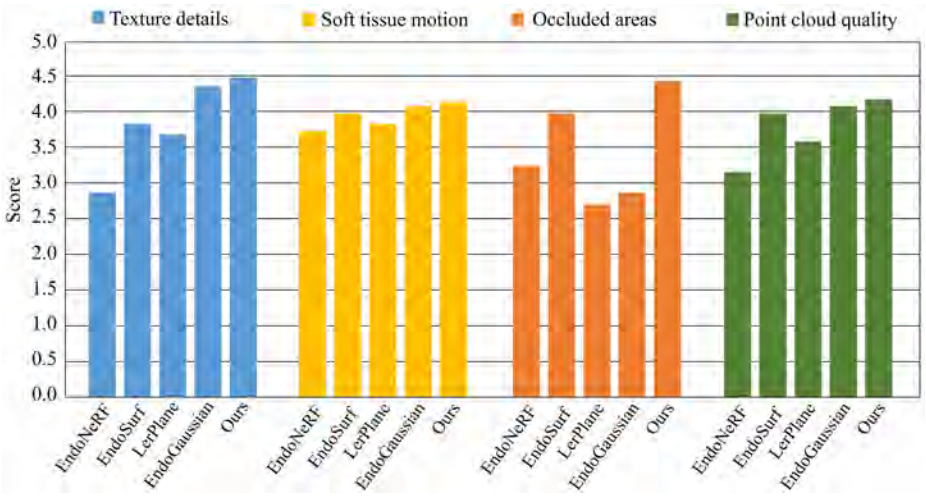

Fig. 2. Average scores (1-5) of the video quality and point cloud quality in user study.

For a fair comparison, we set the video frame rate of each method to a fixed value of 10 FPS when assessing video quality. The results (Fig. 2) indicate that our method significantly outperforms the other methods in terms of video quality, as it more naturally reconstructs occluded areas while maintaining high-quality reconstruction of visible regions. Additionally, due to the depth information of the stereo video is used as supervision in the reconstruction process, the point cloud quality of all methods is relatively close.