Efficient Data-driven Scene Simulation using Robotic Surgery Videos via Physics-embedded 3D Gaussians

Zhenya Yang, Kai Chen, Yonghao Long, and Qi Dou[✉]

The Chinese University of Hong Kong, Hong Kong SAR, China
`qidou@cuhk.edu.hk`

Abstract. Surgical scene simulation plays a crucial role in surgical education and simulator-based robot learning. Traditional approaches for creating these environments with surgical scene involve a labor-intensive process where designers hand-craft tissues models with textures and geometries for soft body simulations. This manual approach is not only time-consuming but also limited in the scalability and realism. In contrast, data-driven simulation offers a compelling alternative. It has the potential to automatically reconstruct 3D surgical scenes from real-world surgical video data, followed by the application of soft body physics. This area, however, is relatively uncharted. In our research, we introduce 3D Gaussian as a learnable representation for surgical scene, which is learned from stereo endoscopic video. To prevent over-fitting and ensure the geometrical correctness of these scenes, we incorporate depth supervision and anisotropy regularization into the Gaussian learning process. Furthermore, we apply the Material Point Method, which is integrated with physical properties, to the 3D Gaussians to achieve realistic scene deformations. Our method was evaluated on our collected in-house and public surgical videos datasets. Results show that it can reconstruct and simulate surgical scenes from endoscopic videos efficiently—taking only a few minutes to reconstruct the surgical scene-and produce both visually and physically plausible deformations at a speed approaching real-time. The results demonstrate great potential of our proposed method to enhance the efficiency and variety of simulations available for surgical education and robot learning.

Keywords: Soft Tissue Simulation · Gaussian Splatting · Surgical Video

1 Introduction

Endoscopic scene simulation is fundamental for surgical training, education and learning-based surgical robot automation [6, 7, 14]. Despite much efforts [1, 3, 4] to simulate deformable objects in anatomical scenes, current solutions rely on manually designed textures that are time-consuming and not scalable. These textures often fail to capture the realistic appearance of various tissues and endoscopic illuminations in real data. Recent advancements in generative AI and

3D reconstruction techniques [5, 13, 22] have raised interest in developing an efficient data-driven surgical scene simulation pipeline, i.e., *can we automatically generate photo-realistic and interactive scenes by only using surgical video*. One feasible way could be to first reconstruct 3D representation of a surgical scene from stereo video, then integrate physics into this scene representation for physically-based simulation. However, achieving this goal is challenging.

Surgical scene 3D reconstruction has been studied in recent works such as EndoNeRF [25], EndoSurf [31], LerPlane [29] and ForPlane [28]. These NeRF-based methods can be used to derive simulatable representation, but implementing this typically requires sophisticated post-processing [24,26] for tetrahedral mesh convert and shape representation transition [15, 18]. Therefore, using NeRF-based reconstruction results for simulation is suboptimal in practice. Very recently, Gaussian Splatting (GS) [9] has emerged as a promising alternative to NeRF [17], offering superior 3D reconstruction results and faster inference speed overall. Several concurrent works [8,12,32] have applied 3D-GS for surgical reconstruction. Unlike NeRF, 3D-GS uses explicit Gaussian representation, making it more suitable for shape editing [2], deformation tracking [16] and simulation [27]. Notably, a successful simulation from 3D-GS relies on clean Gaussians generated from the data, which requires the GS-based reconstruction framework to be particularly designed. For instance, EndoGS [32] is not suitable for simulation purpose, because it tends to generate floating or slim Gaussians in order to fit images, which would lead to noticeable artifacts in simulations (see Fig. 3 in Sec. 3).

Meanwhile, simulating soft-body has been studied via various methods [19, 20,23]. Among them, Material Point Method (MPM) [23] is well-suited for endoscopic scene simulation due to its realism and efficiency. MPM is a physically based simulation method that combines particles and grids to accurately model the deformable objects. A prior work PhysGaussian [27] has promisingly shown the feasibility to simulate 3D Gaussians as deformable objects integrating MPM. The well-defined nature of material point makes MPM suitable to simulate 3D Gaussian while the learned appearance features of 3D Gaussian enhance the visual realism of MPM. Due to the compatibility between MPM and 3D Gaussian, the physics could be naturally embedded into 3D Gaussians by updating them using MPM. Inspired by PhysGaussian, we hope to represent the endoscopic scene with 3D Gaussian and perform simulation on this representation. However, due to the limited movement range of camera in narrow endoscopic space, the 3D Gaussians trained from surgical videos are prone to overfitting which may introduce artifacts in simulation. Overcoming this problem is important for performing surgical scene simulation in a fully automated pipeline.

In this paper, we endeavor to simulate reconstructed surgical scenes captured from single-viewpoint stereo endoscope in a completely data-driven manner (called SimEndoGS). We embark on integrating the emerging 3D Gaussian Splatting and MPM framework to tackle surgical scene simulation. We summarize our contributions as follows: **1)** An efficient and automated pipeline consisting of designed 3D-GS reconstruction module and a subsequent efficient MPM simulation module. **2)** A geometric regularization is proposed to overcome overfitting issue
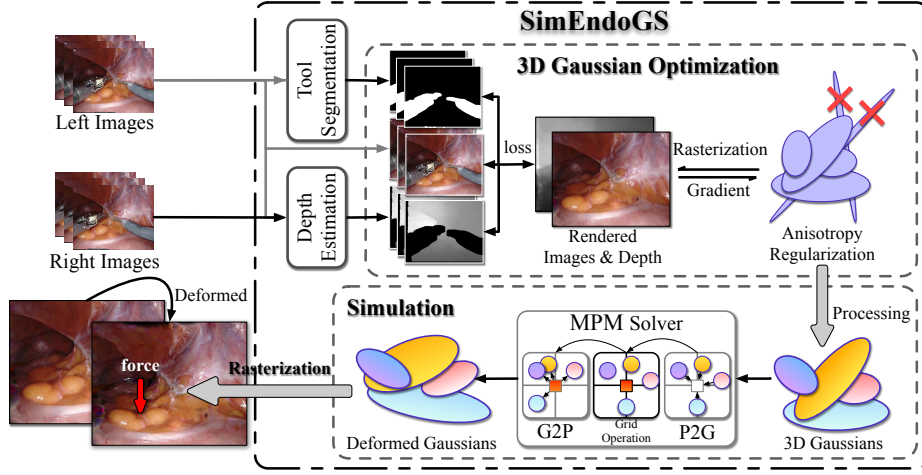
Fig. 1. An overview of the proposed data-driven surgical simulation framework. It consists of automatic scene reconstruction and physically-based scene simulation using 3D Gaussians.

of Gaussian Splatting on endoscopic data. **3)** The Neo-Hookean model and an adapted MPM solver for 3D Gaussian are seamlessly integrated into our pipeline to perform physics-embedded soft tissue simulation. Experimental results on our robotic surgery videos demonstrate the success of our data-driven reconstruction and simulation pipeline, which can support soft-tissue interactions with high fidelity and approaching real-time speed.

2 Method

2.1 Overview of the Data-driven Scene Simulation Framework

Let $\mathcal{V} = \{(\boldsymbol{I}_i^l, \boldsymbol{I}_i^r)\}_{i=1}^T$ be a stereo surgical video with T frames, we aim to develop a fully automatic framework to reconstruct the simulation environment from the video and perform physically-based endoscopic tissue simulation with high fidelity. First of all, we resort to the recent Gaussian Splatting technique for efficient surgical scene reconstruction in Section 2.2. It represents the surgical scene with a group of 3D Gaussians. Similar to [25, 28, 32], we leverage a segmentation model to localize tool regions for each video frame. The tool mask is used for occlusion-free surgical scene reconstruction. Conventional GS-based reconstruction is optimized by minimizing the image-level difference between the reconstructed scene and the original video frame. When it is applied to highly dynamic surgical scenes with limited camera movement range, it is inclined to use floating or slim Gaussians to fit high-frequency image details. Unfortunately, this overfitted representation suffers severe artifacts during simulation. To tackle this problem, we present a novel geometrical regularization method in Section

2.3, which leverages the stereo depth guidance and anisotropy regularization to obtain geometrically plausible scenes suitable for simulation. Based on the improved GS-based surgical scene representation, we develop MPM-based simulation in Section 2.4 for physically based soft tissue simulation. Fig. 1 depicts an overview of our proposed simulation method.

2.2   3D Gaussians for Endoscopic Scene Representation

We represent the endoscopic scene as a group of 3D Gaussians. A 3D Gaussian could be parameterized as $(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{c}, \boldsymbol{\sigma})$, which are corresponding to the position, covariance matrix, color and opacity respectively. These features enable a Gaussian to effectively represent a small area of endoscopic tissue. The 3D Gaussian is optimized using gradients backpropagated from the rendering loss. To render the endoscopic scene composed of 3D Gaussians from a specific view, the covariance matrix of Gaussian can be projected at first: $\hat{\boldsymbol{\Sigma}} = \mathbf{J}\mathbf{W}\boldsymbol{\Sigma}\mathbf{W}^T\mathbf{J}^T$, where $\mathbf{W}$ and $\mathbf{J}$ are matrices related to view transformation and projection, $\hat{\boldsymbol{\Sigma}}$ is the projected covariance matrix which is used to determine whether a pixel is influenced by current Gaussian. After projection, the value of each pixel is computed as a weighted sum of the color features:

$$C(\mathbf{p}) = \sum_{i \in N} \boldsymbol{c_i}\alpha_i \prod_{j=1}^{i-1}(1 - \alpha_j), \tag{1}$$

where N is the number of Gaussians influencing the shading of pixel $\mathbf{p}$ and $\alpha_i = \sigma_i\exp(-\frac{1}{2}(\mathbf{p} - \boldsymbol{\mu_i})^T\boldsymbol{\Sigma'}_i^{-1}(\mathbf{p} - \boldsymbol{\mu_i}))$. To reconstruct an endoscopic scene without surgical tool occlusion, we apply masked L1 loss function at pixel $\mathbf{p}$ utilizing the k-th ground truth image $\boldsymbol{I}_k$, the corresponding tool mask $\boldsymbol{M}_k$ and the rendering result $C(\mathbf{p})$:

$$\mathcal{L}_{color}(\mathbf{p}) = (1 - \boldsymbol{M}_k(\mathbf{p})) * |C(\mathbf{p}) - \boldsymbol{I}_k(\mathbf{p})|. \tag{2}$$

This loss function allows us to effectively utilize surgical video information, as areas occluded by instruments in one frame may be visible in other frames.

2.3   Geometrically Regularized Optimization for 3D Gaussian

In endoscopic scene, the limited movement range of camera makes it difficult for Gaussian Splatting to learn the geometry information and then lead to overfitting. To overcome this problem, we initialize the positions of Gaussians by reprojecting the depth maps $\{\boldsymbol{D}_i\}_{i=1}^T$ to get a reasonable geometric structure before the training. During the optimization, we add the Huber loss between estimated depth maps and rendered depth maps to the objective function of original Gaussian Splatting framework. The depth value of 3D Gaussian at pixel $\mathbf{p}$ shares the same computation pattern with Eq. 1. Our complete loss function is:

$$\mathcal{L}(\mathbf{p}) = \mathcal{L}_{color}(\mathbf{p}) + \eta\mathcal{L}_{Huber}(\boldsymbol{D}_k(\mathbf{p}), \sum_{i \in N} d_i\alpha_i \prod_{j=1}^{i-1}(1 - \alpha_j)) \tag{3}$$

where d_i is the z value of the i-th Gaussian, $\boldsymbol{D}_k$ is the k-th estimated depth map and hyperparameter η is used to control the strength of depth regularization. The depth regularization enforces Gaussians to distribute near the tissue surface and penalize the generation of floater to reduce artifact. The artifacts in simulation results can also be caused by the Gaussians with slender shape. Therefore, we perform anisotropy regularization to prune Gaussians with very slim shapes during the optimization. The scale ratio is used to determine whether a Gaussian should be pruned: $Ratio(p) = \max(\mathbf{S_p})/\min(\mathbf{S_p})$, where $\mathbf{S_p}$ is the scaling tensor of Gaussian p. We prune Gaussians whose scale ratio surpasses the predefined threshold γ during the training.

2.4   Physically-based 3D Gaussians Simulation

Before simulation, we perform the Gaussian padding for better simulation effect. Directly simulating Gaussians without padding often leads to surface crash as shown in Fig. 4. We compute the opacity field O using the following equation on a uniform $100 \times 100 \times 100$ Eulerian grid: $O(\boldsymbol{x}) = \sum_i \sigma_i \exp(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x_i})^T \mathbf{\Sigma_i}^{-1}(\boldsymbol{x} - \boldsymbol{x_i}))$, where $\boldsymbol{x}$ is the position of grid node and $\boldsymbol{x_i}$ is the center of Gaussian i surrounding the grid node. If the value of current node in opacity field is less than that of nodes closer to camera, it indicates that the current grid node is behind the tissue surface and we will pad a Gaussian at this place.

We integrate the Material Point Method into our framework to perform physically based tissue simulation on reconstructed scenes represented in 3D Gaussians. The position of 3D Gaussians can be directly updated as lagrangian particles in MPM. To let 3D Gaussians capture the deformation of material, [27] propose to update the covariance matrix of Gaussian as follows:

$$\mathbf{\Sigma}' = \mathbf{F}\mathbf{\Sigma}\mathbf{F}^T, \tag{4}$$

where $\mathbf{F}$ is the deformation gradient obtained from MPM solver, $\mathbf{\Sigma}$ is the initial covariance matrix and $\mathbf{\Sigma}'$ is the updated covariance matrix. We utilize Neo-Hookean [21] constitutive model in MPM to predict tissue deformation because of its simplicity and computation efficiency. The first Piola-Kirchoff stress of Neo-Hookean model, denoted as $\mathbf{PK_1}$, is computed as follows:

$$\mathbf{PK_1} = \mu(\mathbf{F} - \mathbf{F^T}) + \lambda \log(J)\mathbf{F^{-T}}, \tag{5}$$

where μ and λ are lame parameters computed using Young's modulus E and Poisson ratio ν with following equations:

$$\mu = \frac{E}{2(1+\nu)}, \ \lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}. \tag{6}$$

By adjusting the values of Young's modulus E and Poisson ratio ν, we can control the physical behavior of the endoscopic tissue, such as its stiffness, in a physically interpretable way.

3 Experiments

We extensively evaluated our proposed data-driven surgical simulation on five different endoscopic scenes, using corresponding stereo videos collected from the publicly available datasets [25, 30] or our in-house data. These videos encompassed common surgical procedures involving the manipulation, pushing and pulling of deformable tissues. We compared our method with an GS-based baseline. Similar to [8, 12, 32], we quantitatively evaluated the surgical scene reconstruction using PSNR. For the simulation performance, since it is difficult for a quantitative evaluation, we reported the simulation efficiency and qualitatively measured the simulation performance by comparing the texture and geometry details under different simulation interactions. We also conducted an ablation study to evaluate the proposed depth supervision module, anisotropy regularization module and Gaussian padding module through qualitative comparison.

3.1 Implementation Details

In our implementation, we set the delta in Huber loss to 0.2, the weight of depth loss $\eta = 0.3$ and the anisotropy regularization $\gamma = 10$. All endoscopic scenes used in our experiments are optimized with 7000 iterations. We utilize STTR-light [11] pretrained on Scene Flow to estimate stereo depth maps of binocular surgical video frames. The tool masks are automatically generated by SegmentAnything [10]. For simulation, we perform 80 substeps in each step and the timestep of each substep is 0.0005s. All experiments were conducted on a computer equipped with an Intel(R) Xeon(R) W-2223 CPU and RTX3090 GPU.

3.2 Qualitative and Quantitative Results

We evaluated the proposed data-driven surgical scene simulation on various surgical scenarios. Fig. 2 presents the qualitative evaluation results. Given the stereo surgical video, in which the tissue would be frequently occluded by the tool due to the surgical operation, our method can robustly reconstruct occlusion-free tissues with realistic textures. It is the basis for achieving high-fidelity surgical simulation. To assess the quality of the simulation, we simulated interactive actions with the reconstructed simulation environment by applying forces in various directions at different positions. As shown in Fig. 2, a closer examination of the highlighted areas reveals that our simulation method not only generates visually plausible shape deformations but also preserves the visual consistency of the texture. Table 1 further reports the detailed processing time and simulation efficiency for each endoscopic scene. Without heavy manual adjustments, our method is highly scalable and is able to efficiently consume various surgical videos for high-quality surgical simulation. We refer readers to the supplementary video for more simulation results.

To further demonstrate the advantage of our method, we compared it with the EndoGS baseline [32]. As shown in Table 1, we observed that EndoGS usually achieves a higher PSNR than ours, which indicates that its reconstructed surgical
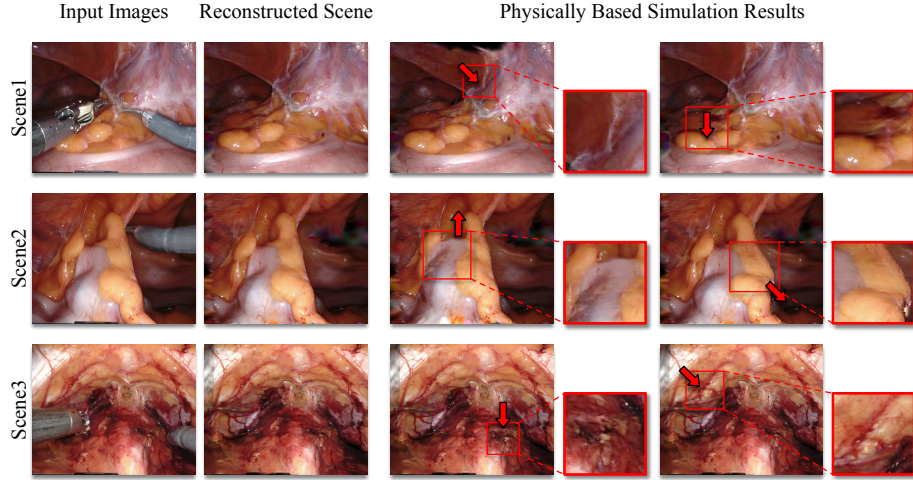
Input Images   Reconstructed Scene   Physically Based Simulation Results

Fig. 2. Qualitative evaluation of simulation performance. The external forces are indicated using red arrows. The corresponding simulation videos are included in the supplementary material.

Table 1. Quantitative results (standard deviation in parentheses).

Scenes	Training	Ours PSNR	EndoGS PSNR	Processing	Gaussian Number	Simulation FPS
Scene1	65.0(0.8)s	34.77(0.04)	40.44(0.30)	0.423(0.005)s	78,712	20
Scene2	61.3(0.5)s	37.82(0.02)	39.94(0.24)	0.495(0.006)s	74,484	21
Scene3	60.7(0.5)s	36.05(0.24)	36.28(0.04)	0.453(0.002)s	58,223	25

scene is closer to the original video frame. We argue that it is because EndoGS is inclined to fit the high-frequency scene content (e.g., specular light indicated by the white dashed box in Fig. 3) with floater. As shown in Fig. 3, the reconstructed scene would result in obvious artifacts when it is applied to simulation. In contrast, our method use geometric regularization to obtain scenes with more reasonable space structure. It achieves comparable reconstruction quality to EndoGS. More importantly, our method significantly outperforms EndoGS in terms of the simulation quality. It not only maintains reasonable geometric structures and tissue textures but also produces physically realistic tissue deformations, as highlighted in the specific area of Fig. 3. These results fully demonstrate the superiority of our proposed method for endoscopic scene simulation.

To valid each module of our method, we conducted a qualitative ablation study. Fig. 4 presents the experimental results. When anisotropy regularization is omitted, the slender kernels of the Gaussians become exposed under large-scale deformation, resulting in a fur-like artifact. In contrast to our results, simulations without depth supervision exhibit floating artifacts for lack of geometric
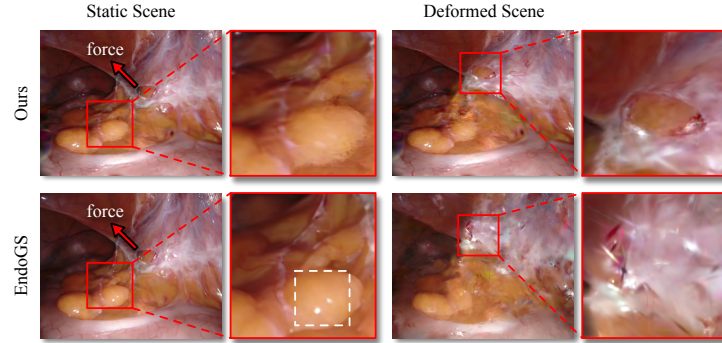
Fig. 3. Comparison with EndoGS [32]. The comparison between ours method and EndoGS on reconstruction and simulation.
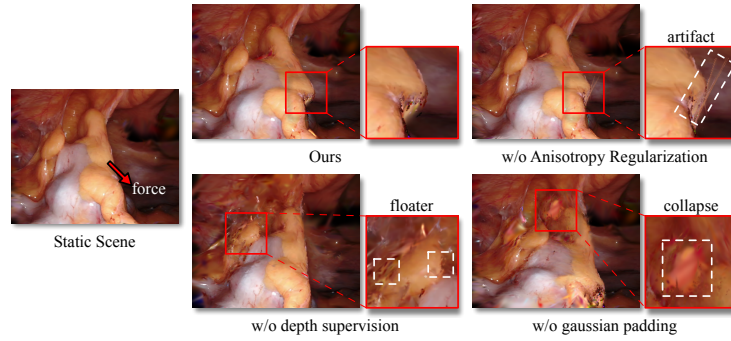

Fig. 4. Ablation study. We compare our base simulation result and simulation results w/o depth supervision, gaussian padding or anisotropy regularization. The artifacts are highlighted using white dashed boxes.

regularization in training, leading to a noisy result after deformation. Simulations without Gaussian padding result in a thin surface that is prone to collapse under external forces. The corresponding result depicted in Fig. 4 exhibits noticeable cracks and exposes the underlying material.

4   Conclusion

This paper introduces a novel framework based on Gaussian Splatting for automated surgical scene reconstruction from stereo surgical videos and physically-based endoscopic scene simulation with user-defined interactions. We coherently utilize the 3D Gaussian representation for reconstruction and simulation for convenient simulation, efficient visualization and realistic visual results. We have specifically designed an optimization strategy to enhance the suitability of our learned 3D Gaussians for subsequent simulation tasks. Our method achieves a

high degree of automation and superior efficiency, seamlessly transforming surgical videos into interactive simulation scenes. We anticipate that the integration of generative AI and 3D reconstruction techniques will inspire the development of interactive and highly realistic surgical scene generation, benefiting surgical training and enhancing the learning capabilities of surgical robots.

References

1. Bar-Meir, S.: Simbionix simulator. Gastrointestinal Endoscopy Clinics **16**(3), 471–478 (2006)
2. Chen, Y., Chen, Z., Zhang, C., Wang, F., Yang, X., Wang, Y., Cai, Z., Yang, L., Liu, H., Lin, G.: Gaussianeditor: Swift and controllable 3d editing with gaussian splatting. arXiv preprint arXiv:2311.14521 (2023)
3. Faure, F., Duriez, C., Delingette, H., Allard, J., Gilles, B., Marchesseau, S., Talbot, H., Courtecuisse, H., Bousquet, G., Peterlik, I., et al.: Sofa: A multi-model framework for interactive physical simulation. Soft tissue biomechanical modeling for computer assisted surgery pp. 283–321 (2012)
4. Hirota, G., Fisher, S., State, A.: An improved finite-element contact model for anatomical simulations. The Visual Computer **19**, 291–309 (2003)
5. Hong, Y., Zhang, K., Gu, J., Bi, S., Zhou, Y., Liu, D., Liu, F., Sunkavalli, K., Bui, T., Tan, H.: Lrm: Large reconstruction model for single image to 3d. arXiv preprint arXiv:2311.04400 (2023)
6. Huang, T., Chen, K., Li, B., Liu, Y.H., Dou, Q.: Demonstration-guided reinforcement learning with efficient exploration for task automation of surgical robot. arXiv preprint arXiv:2302.09772 (2023)
7. Huang, T., Chen, K., Wei, W., Li, J., Long, Y., Dou, Q.: Value-informed skill chaining for policy learning of long-horizon tasks with surgical robot. arXiv preprint arXiv:2307.16503 (2023)
8. Huang, Y., Cui, B., Bai, L., Guo, Z., Xu, M., Ren, H.: Endo-4dgs: Endoscopic monocular scene reconstruction with 4d gaussian splatting. arXiv preprint arXiv:2401.16416 (2024)
9. Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G.: 3d gaussian splatting for real-time radiance field rendering. ACM Transactions on Graphics **42**(4) (2023)
10. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Dollár, P., Girshick, R.: Segment anything. arXiv preprint arXiv:2304.02643 (2023)
11. Li, Z., Liu, X., Drenkow, N., Ding, A., Creighton, F.X., Taylor, R.H., Unberath, M.: Revisiting stereo depth estimation from a sequence-to-sequence perspective with transformers. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp. 6197–6206 (October 2021)
12. Liu, Y., Li, C., Yang, C., Yuan, Y.: Endogaussian: Real-time gaussian splatting for dynamic endoscopic scene reconstruction. arXiv preprint arXiv:2401.12561 (2024)
13. Liu, Y., Zhang, K., Li, Y., Yan, Z., Gao, C., Chen, R., Yuan, Z., Huang, Y., Sun, H., Gao, J., He, L., Sun, L.: Sora: A review on background, technology, limitations, and opportunities of large vision models. arXiv preprint arXiv:2402.17177 (2024)
14. Long, Y., Wei, W., Huang, T., Wang, Y., Dou, Q.: Human-in-the-loop embodied intelligence with interactive simulation environment for surgical robot learning. arXiv preprint arXiv:2301.00452 (2023)

15. Lorensen, W.E., Cline, H.E.: Marching cubes: A high resolution 3d surface construction algorithm. In: Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques. p. 163–169. SIGGRAPH '87, Association for Computing Machinery, New York, NY, USA (1987). https://doi.org/10.1145/37401.37422, https://doi.org/10.1145/37401.37422
16. Luiten, J., Kopanas, G., Leibe, B., Ramanan, D.: Dynamic 3d gaussians: Tracking by persistent dynamic view synthesis. In: 3DV (2024)
17. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM **65**(1), 99–106 (2021)
18. Molino, N., Bridson, R., Fedkiw, R.: Tetrahedral mesh generation for deformable bodies. In: Proc. Symposium on Computer Animation. vol. 8 (2003)
19. Müller, M., Heidelberger, B., Hennix, M., Ratcliff, J.: Position based dynamics. Journal of Visual Communication and Image Representation **18**(2), 109–118 (2007)
20. Reddy, J.N.: Introduction to the finite element method. McGraw-Hill Education (2019)
21. Rivlin, R.S., Rideal, E.K.: Large elastic deformations of isotropic materials iv. further developments of the general theory. Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences **241**(835), 379–397 (1948). https://doi.org/10.1098/rsta.1948.0024, https://royalsocietypublishing.org/doi/abs/10.1098/rsta.1948.0024
22. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684–10695 (2022)
23. Sulsky, D., Zhou, S.J., Schreyer, H.L.: Application of a particle-in-cell method to solid mechanics. Computer Physics Communications **87**(1), 236–252 (1995). https://doi.org/https://doi.org/10.1016/0010-4655(94)00170-7, https://www.sciencedirect.com/science/article/pii/0010465594001707, particle Simulation Methods
24. Tang, J., Zhou, H., Chen, X., Hu, T., Ding, E., Wang, J., Zeng, G.: Delicate textured mesh recovery from nerf via adaptive surface refinement. arXiv preprint arXiv:2303.02091 (2023)
25. Wang, Y., Long, Y., Fan, S.H., Dou, Q.: Neural rendering for stereo 3d reconstruction of deformable tissues in robotic surgery. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 431–441. Springer (2022)
26. Wei, X., Xiang, F., Bi, S., Chen, A., Sunkavalli, K., Xu, Z., Su, H.: Neumanifold: Neural watertight manifold reconstruction with efficient and high-quality rendering support. arXiv preprint arXiv:2305.17134 (2023)
27. Xie, T., Zong, Z., Qiu, Y., Li, X., Feng, Y., Yang, Y., Jiang, C.: Physgaussian: Physics-integrated 3d gaussians for generative dynamics. arXiv preprint arXiv:2311.12198 (2023)
28. Yang, C., Wang, K., Wang, Y., Dou, Q., Yang, X., Shen, W.: Efficient deformable tissue reconstruction via orthogonal neural plane. arXiv preprint arXiv:2312.15253 (2023)
29. Yang, C., Wang, K., Wang, Y., Yang, X., Shen, W.: Neural lerplane representations for fast 4d reconstruction of deformable tissues. arXiv preprint arXiv:2305.19906 (2023)
30. Ye, M., Johns, E., Handa, A., Zhang, L., Pratt, P., Yang, G.Z.: Self-supervised siamese learning on stereo image pairs for depth estimation in robotic surgery. arXiv preprint arXiv:1705.08260 (2017)

31. Zha, R., Cheng, X., Li, H., Harandi, M., Ge, Z.: Endosurf: Neural surface reconstruction of deformable tissues with stereo endoscope videos. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 13–23. Springer (2023)
32. Zhu, L., Wang, Z., Cui, J., Jin, Z., Lin, G., Yu, L.: Endogs: Deformable endoscopic tissues reconstruction with gaussian splatting. arXiv preprint arXiv:2401.11535 (2024)