Reducing Overfitting in 3D Gaussian Splatting using Depth Supervision

T.H.B. Spanhoff
t.h.b.spanhoff@student.tudelft.nl

Supervisor: Xucong Zhang
June 2024

CSE3000 Research Project
Bsc. Computer Science and Engineering TU Delft

3D Gaussian Splatting (3DGS) [2]

- A method for representing 3D scenes using a point cloud of 3D Gaussians. [2]
- Trained by a set of images of a scene
- Useful for media, navigation systems, and medical applications [3, 4]
- Prone to overfitting
  - Caused by limited viewpoint diversity **(Figure 1)**
  - Results in artifacts like floating Gaussians **(Figure 2)**

Figure 1: The original RGB-D images are taken from similar positions that approximate a circle.

Figure 2: The white Gaussians representing the sky look decent on one view, but can be seen floating in mid-air when rendering the scene from a novel view.

Figure 3: Similarly to original 3DGS, RGB images of an object are used to train a Gaussian Splat.

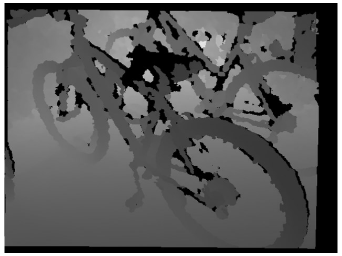

Figure 4: 3DGSw/Depth is trained with images from RGB-D cameras, which also take a depth map as input.

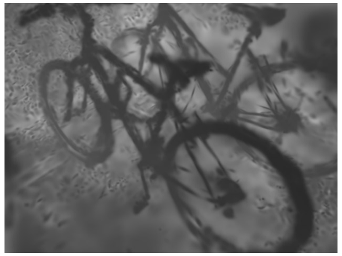

Figure 5: 3DGSw/Depth renders a depth map from a Gaussian Splat by splatting each Gaussian's depth to 2D.

Figure 6: The depth loss is the absolute difference after preprocessing both depth maps.

Research Question

How to exploit additional depth information from RGB-D measurements in 3D Gaussian splatting?

Our Contribution

- Introduction of 3D Gaussian Splatting with Depth (3DGSw/Depth), which incorporates depth supervision from RGBD cameras into the training process.
- Evaluation of 3DGSw/Depth by:
  - Analyzing the loss functions during training
  - Comparing novel views from 3DGS and 3DGSw/Depth's renders

3DGS with Depth

- Trains a Gaussian Splat using gradient descent in a similar way as original 3DGS [2], but uses an additional depth loss factor as follows:

1. Takes an additional Depth Map as input, taken from an RGB-D camera **(Figure 4)**
2. Renders a depth map from the Gaussian Splat by splatting each Gaussian's depth to 2D **(Figure 5)**
3. RGB-D images do not have depth information for every pixel, so data-less pixels are ignored in the depth loss function
4. Rendered and ground truth depth maps can be off by some scaling factor, so they are both normalized
5. The depth loss is the mean of the absolute difference between the ground truth and rendered depth maps. **(Figure 6)**
6. The final loss is a weighted average of the original 3DGS loss [2] on RGB images **(Figure 3)** and the depth loss.

Results

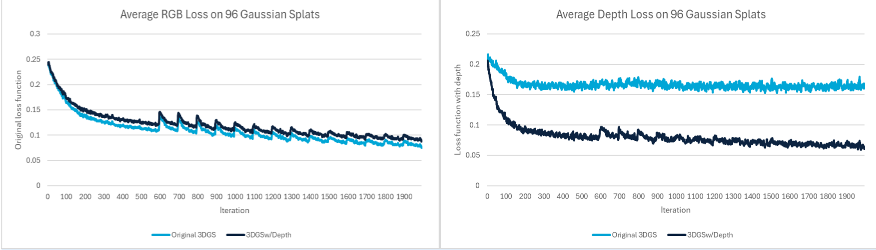

Figure 7: Loss on RGB images (left) and depth maps (right) during training of 192 Gaussian Splats on 96 objects with depth weights 0.0 and 0.5. The spikes that occur with intervals of 100 iterations are caused by the splitting and cloning of Gaussians during 3DGS

Figure 8: No floating artifacts visible after training a model with 3DGSw/Depth

- 96 Gaussian Splats each were trained with original 3DGS and 3DGSw/Depth
- RGB-D object scans from the Redwood dataset [1] were used
- Object scans were subsampled to 200 images
- Splats were trained for 2000 iterations, were 3DGSw/Depth used a depth weight of 0.5
- The depth loss was reduced by a factor of three while the RGB loss increased only slightly **(Figure 7)**
- Models trained with 3DGSw/Depth generally contained less floating artifacts than those trained with original 3DGS **(Figure 2 & 8)**

Conclusion

- Additional depth information from RGB-D measurements can be exploited in 3DGS by rendering and comparing depth maps in a similar fashion as RGB images are compared in 3DGS
- 3DGSw/Depth drastically reduces the depth loss without substantially decreasing the RGB loss
- 3DGSw/Depth reduces floating Gaussians, especially in outdoor scenes.
- For depth information that was incorrect or not synchronized with the RGB images, 3DGSw/Depth can have a huge negative impact on the model's quality.

References

[1] Sungjoon Choi, Qian-Yi Zhou, Stephen Miller, and Vladlen Koltun. A large dataset of object scans. *arXiv:1602.02481*, 2016.

[2] Bernhard Kerbl, Georgios Kopanas, Thomas Leimk¨uhler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4), July 2023

[3] Emmanouil Nikolakakis, Utkarsh Gupta, Jonathan Vengosh, Justin Bui, and Razvan Marinescu. Gaspct: Gaussian splatting for novel ct projection view synthesis, 2024

[4] Matthew Tancik, Ethan Weber, Evonne Ng, Ruilong Li, Brent Yi, Terrance Wang, Alexander Kristoffersen, Jake Austin, Kamyar Salahi, Abhik Ahuja, David Mcallister, Justin Kerr, and Angjoo Kanazawa. Nerfstudio: A modular framework for neural radiance field development. In *Special Interest Group on Computer Graphics and Inteactive Techniques Conference Conference Proceedings*, SIGGRAPH '23. ACM, July 2023.