Improving Gaussian Splatting using image preprocessing
A research documentation of the Advanced Medical Machine Learning Seminar from 2024

Valentin Teutschbein[1]

Digital Health & Machine Learning
Hasso Plattner Institute for Digital Engineering
mail@valteu.com

1 Introduction

3D-Reconstruction technology is widely applied in many areas. For example, the creation of humanoid 3D avatars can be used for medical applications such as the visualization of weight transformation or supporting physical training. These avatars are commonly created using the Gaussian Splatting 3D reconstruction method. In this paper, multiple approaches to enhancing Gaussian Splatting and point cloud creation using image preprocessing are documented, summing up the experiments we conducted as part of the 'Advanced Medical Machine Learning Seminar' at the Hasso-Plattner-Institute. 3D-Reconstruction using Gaussian Splatting usually uses an image dataset of a scene to reconstruct, extracts a 3D colored point cloud of the images using Colmap and then uses this point cloud and the images of the dataset to iteratively improve the 3D Gaussians located around each point as a mean value to better fit the images of the dataset. In this research, we focused on adding one additional step to this pipeline: Preprocessing the images first before extracting the point cloud. We hope that this allows us to enhance the structure of the resulting 3D-Reconstruction, especially for datasets of lower quality images in terms of lightning and sharpness. Additionally, we experimented with different Colmap [5] point cloud extraction configurations and computed metrics for the resulting point clouds.

2 Implementation and execution

In this section, the different ideas, solution approaches and their results are presented. We created the GitHub project 'Point-Cloud-Enhancer'[1] to sum up our work and experiments into one data pipeline. This project includes dataset preprocessing, point cloud extraction, the applied point cloud metric and Gaussian

[1]https://github.com/valteu/point-cloud-enhancer.git

Splatting itself. In the appendix, other failed or not further proceeded approaches are shared as well.

2.1 Initial preprocessing results

For the four image datasets 'train', 'truck', 'drjohnson' and 'playroom' (view in appendix) provided by the Gaussian Splatting project, we used two different methods of image preprocessing: Adaptive-Chromaticity enhancement [6] and edge enhancement [3]. Then we created one Colmap dataset and a corresponding Gaussian Splatting [2] preprocessing result for each of these datasets and for each preprocessing technique, including the default images. On these models, we computed the SSIM, PSNR and LPIPS metrics, which compare the rendered images generated from the 3D reconstruction from the camera perspective of the test images with each corresponding test image. The Idea behind this procedure was that the preprocessing could help the model to match 3D points and help correct areas of the images with sparse lighting. However, the resulting metrics showed for each dataset that the non-enhanced images got the best metric scores, so preprocessing did not help to improve these metric results in any of the tested cases.

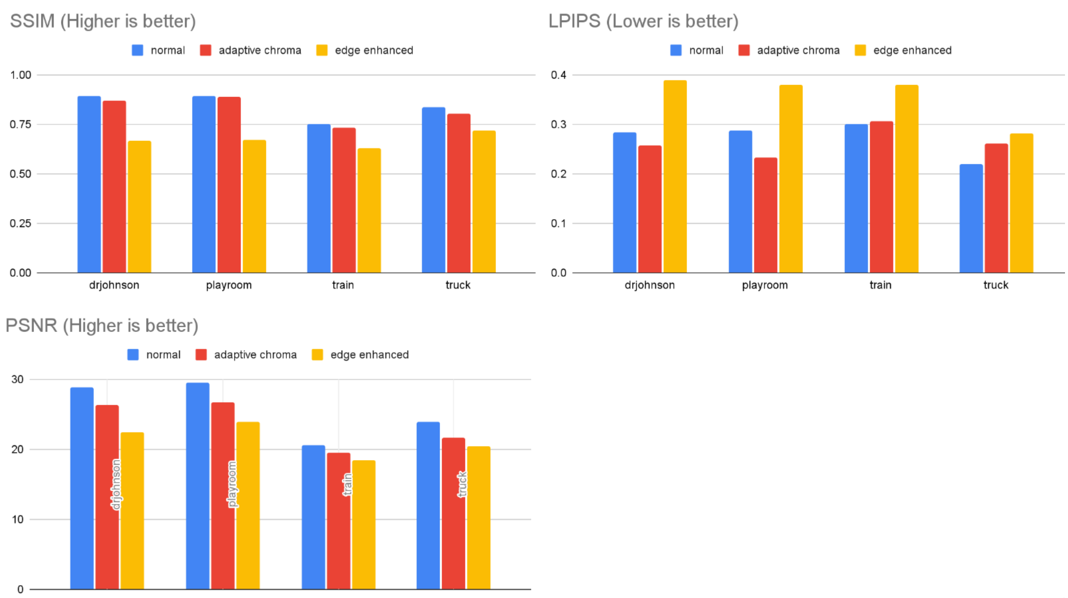

Figure 1: Overview of the 4 datasets in the 3 states of not preprocessed, adaptive chromaticity enhanced and edge enhanced in the 3 image metrics SSIM, LPIPS and PSNR

2.2 Improving structure from motion (SfM)

At next, we figured out that, in many cases, the preprocessed images were sorted out as distorted by Colmap. Thus, we tried using different Colmap matching algorithms. Therefore, we modified three Colmap flags ('–SiftMatching.guided_matching', '–SiftExtraction.estimate_affine_shape' and '–SiftExtraction.domain_size_pooling') which we set true. Additionally, we added four new image datasets of small to medium size (60–150 images) of poor lighting quality to our experiment (view in appendix) to see which influence these preprocessing technologies have on images taken in dark conditions. Besides that, we replaced the adaptive chromaticity enhancement with exposure enhancement [1] and introduced an additive preprocessing method which exposure enhanced the edge enhanced images. However, again, the three metrics (SSIM, PSNR and LPIPS) still showed the not preprocessed images ahead of the preprocessed datasets. The only obvious improvement was that the two darkest bin datasets could not be reconstructed without preprocessing because Colmap did not find sufficient matching 3D points. But when taking a look at the point clouds created by Colmap, we figured that the point clouds of the modified Colmap matching pipeline looked much denser but also had more noise. Thus, we had to critically view our decision about which metrics to use, which is explained in the following section.

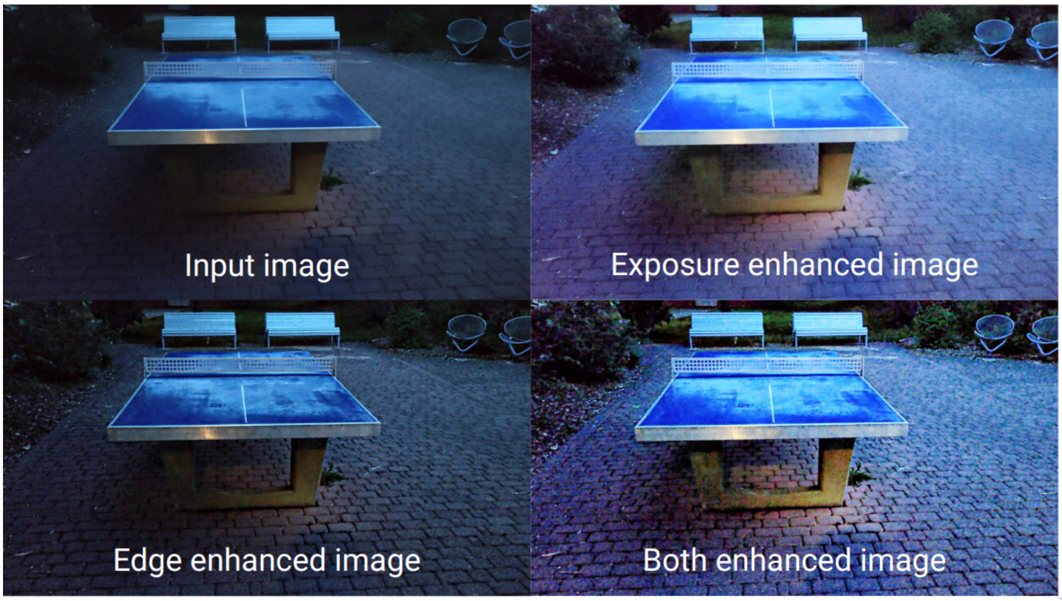

Figure 2: This figure shows the 4 different (non)-preprocessed combinations used in the example of the 'pingpong' dataset

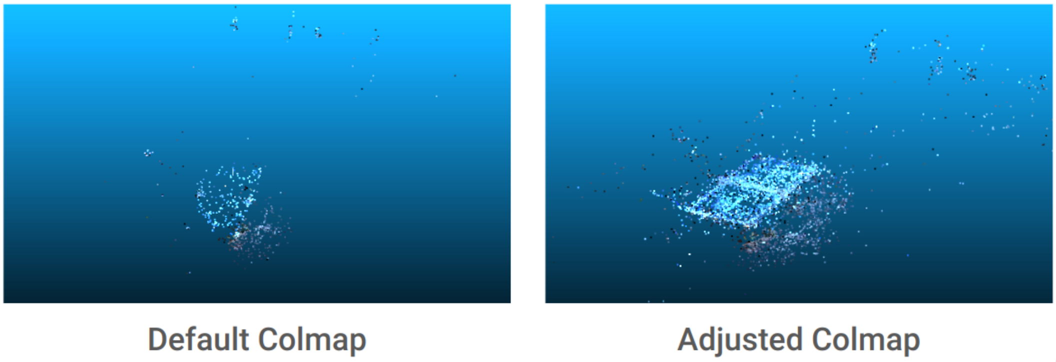

Figure 3: This image shows two point clouds extracted from the 'pingpong' dataset. The left one used the default Colmap modification and the right one used the adjusted one. As you can see, the adjusted point cloud has a higher level of detail and more density

2.3 Metric discussion

Using the SSIM, PSNR and LPIPS metrics to evaluate model quality is suitable for seeing how well test images are to be rendered inside the created 3D model from the corresponding camera perspective. However, we don't want to compare how well preprocessed and default image datasets are reconstructed; instead, we want to see whether preprocessing improves the model quality overall. So we want to measure if preprocessing the images creates a model with better quality compared to the default image model. Therefore, we decided to focus on the point clouds layer of the model pipeline and computing metrics using a support vector regression (SVR) model trained on 64 features of a subjective assessment database provided by the NR-3DQA project [7]. With this metric, we can compare how well the resulting colored point clouds are subjectively perceived, which offers a way of comparing the models to each other.

2.4 Applying new metrics

In the following experiment, the SVR-generated metrics were used to determine the point cloud quality. This method clearly favored the modified Colmap matching approach over the default Colmap matching in median by over 18% (measuring the score difference relative to the resulting score). In just two datasets, this matching metric was lower than the default one. The best measured preprocessing technique was applying both, edge and exposure enhancement. Additionally, when comparing the metric results of the non preprocessed datasets with the datasets with both preprocessing applied, the median improvement lays at over 45% and at least by 12% if ignoring the two datasets, no Colmap could be created when using no preprocessing.

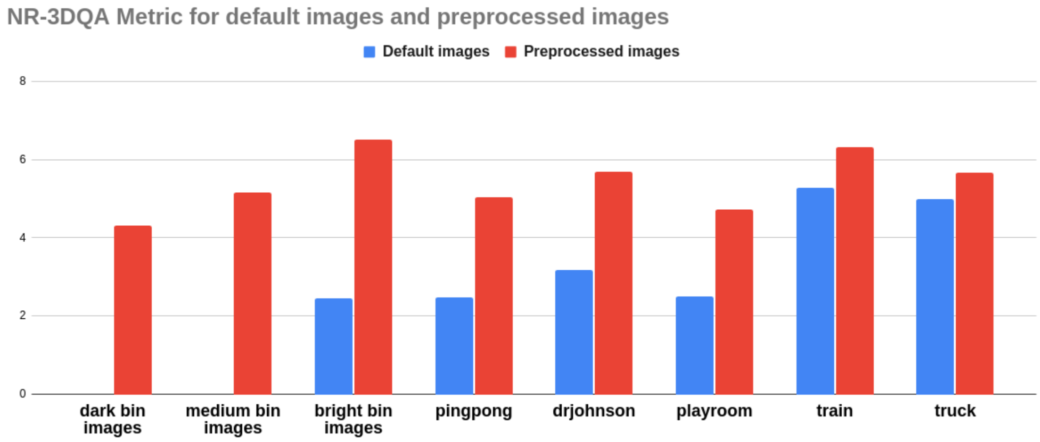

Figure 4: Resulting NR-3DQA metrics for the 8 datasets (Higher is better, scale is from 0 to 10). The blue bars show the metric values of the default Colmap configuration without preprocessing, while in red the metric values for the modified Colmap together with edge enhancement and exposure enhancement are shown

2.5 Large dataset degradation

Since in the last experiment the datasets of poor image quality all had just a comparably small amount of images and using naturally bad quality datasets limits control the over degradation level, we decided to additionally test our preprocessing with the four provided datasets from Gaussian Splatting by degrading them, using three strengths of blur and dark. Because of the huge time expense of creating four variations of preprocessing for these six variations of degradation, we decided to continue just with the 'drjohnson' dataset (view in appendix). There, we compared the results for both Colmap configurations and each preprocessing technique for each degradation. As a result, we compared the default Colmap configuration without preprocessing to the modified Colmap version with both preprocessing technologies applied. Here we have a median metric increase of over 34% and a metric increase of at least over 30%. After computing the metrics, we additionally took a look at the generated point clouds and Gaussian Splatting models. On the point cloud level, the dark datasets were visually enhanced the most with exposure enhancement. However, we also got some mismatches in preprocessed datasets, such that one room in the point cloud got created twice. When taking a look at the generated Gaussian Splatting models, we saw subjectively more artifacts in the preprocessed datasets, which thus appeared less appealing than the default images. We assume this might be due to inconsistent preprocessing within the same dataset because an object might get different preprocessed looks from different images, which might increase the difficulty of creating a consistent 3D representation of it.

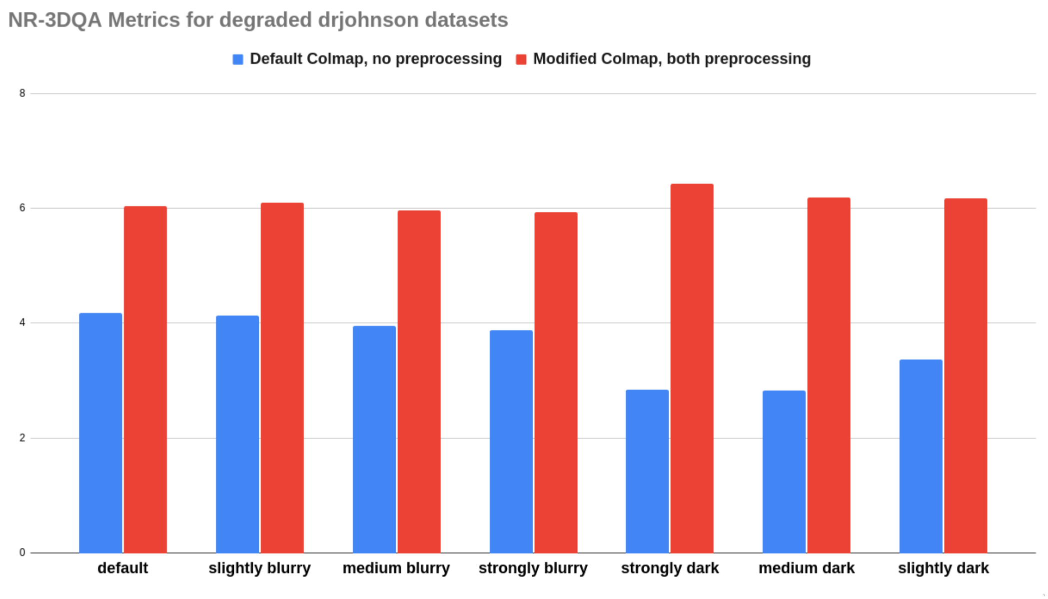

Figure 5: The figure shows how the both preprocessed datasets compare to the not preprocessed datasets in the point cloud metric

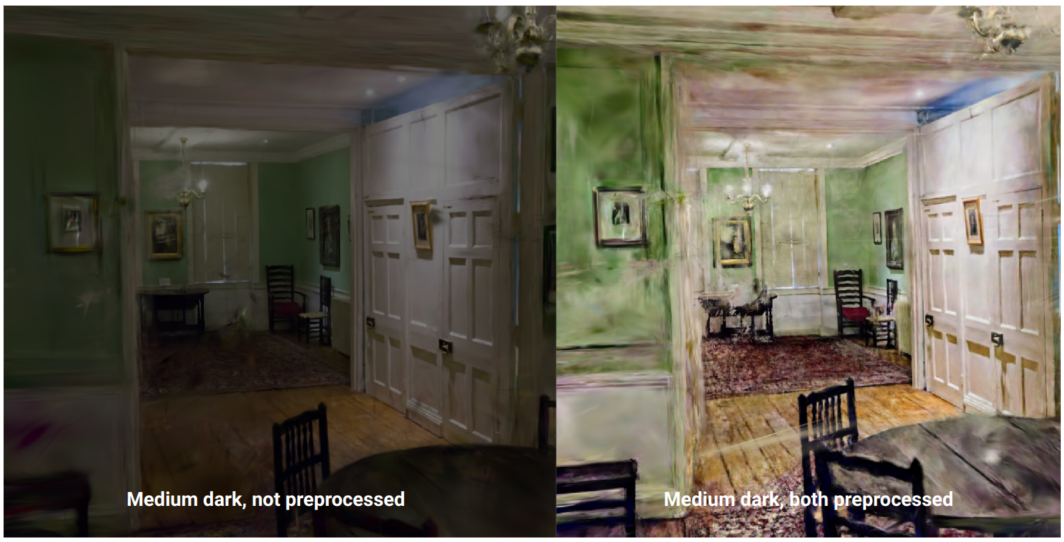

Figure 6: This figure shows the artefacts visible especially in the preprocessed image dataset on the right. Both datasets are based on the medium dark degradation

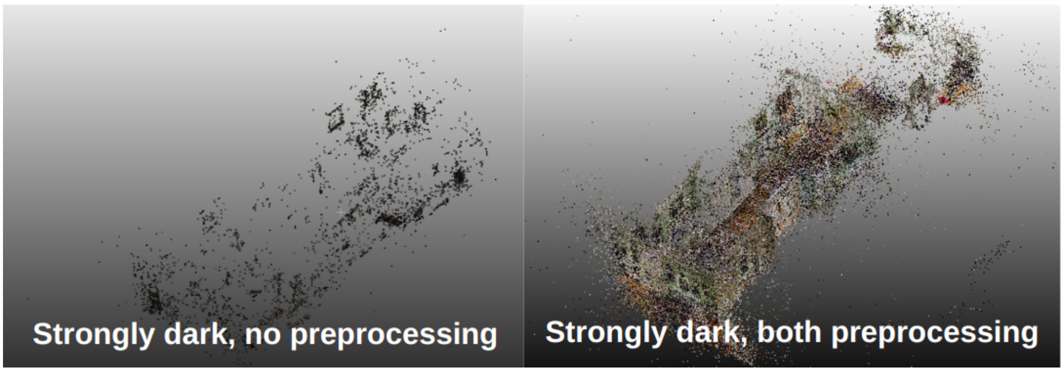

Figure 7: The figure shows the significant improvement of preprocessing on the strongly dark degradated dataset. However the preprocessed version has a mapping issue with one room on the right, which should not exist. Also noise in terms of points outside the scene is visible

3 Conclusion

All in all, in our test setup, after modifying the matching algorithm of Colmap and preprocessing both exposure and edge enhancements, we got much denser point clouds from image datasets, which also had a better structural appearance. This is also represented by the metrics we used. However, this preprocessing method comes with its limits. It appears to be much more noisy, mainly because of the different matching and also the created Gaussian Splatting results visually look worse compared to the non-preprocessed version. Both issues need to be addressed before practically applying this method. Additionally, preprocessing the datasets is another time-consuming step in the 3D reconstruction pipeline.

4 Future Directions

To address our issue of noisy point clouds we propose adding a point cloud post processing step to the pipeline. Additionally the consistency of preprocessing needs to be ensured.

5 Appendix

5.1 Separating structure from color

While experimenting with image preprocessing for enhancing Gaussian Splatting, we initially started with the idea of separating the Colmap point cloud extraction process into two phases: the first being the generation of the 3D point locations for the point cloud itself, and the second one being the correct coloring of each point. We tried creating this using the edge-enhanced images for the structural point cloud creation and the exposure-enhanced images for selecting the colors per point. We also appended this approach using the alpha-blending of both preprocessing techniques. However, we did not further proceed with this approach since, in both cases, we were facing worse SSIM, PSNR and LPIPS results compared to the default Colmap configuration.

5.2 Experimenting with Glomap

During our research project, the Glomap [4] project was released. We tested the impact of replacing Colmap-SfM with Glomap. However, the results were slightly worse or on-par with Colmap, which made us not further proceed in this approach.

5.3 Datasets

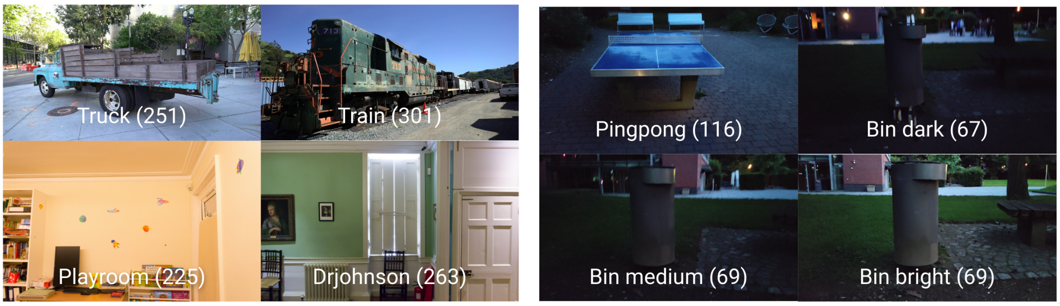

Figure 8: This figure shows example images of all 8 datasets used in the process labeled with the corresponding name and the number of images

Figure 9: This figure shows the 6 degraded images of the drjohnson dataset used

References

[1] K. X. Haoyuan Wang and R. W. Lau. “Local Color Distributions Prior for Image Enhancement”. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. 2022.

[2] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis. “3D Gaussian Splatting for Real-Time Radiance Field Rendering”. In: *ACM Transactions on Graphics* 42.4 (July 2023).

[3] W. Liu, P. Zhang, X. Huang, J. Yang, C. Shen, and I. Reid. “Real-time image smoothing via iterative least squares”. In: *ACM Transactions on Graphics (TOG)* 39.3 (2020), pages 1–24.

[4] L. Pan, D. Barath, M. Pollefeys, and J. L. Schönberger. “Global Structure-from-Motion Revisited”. In: *European Conference on Computer Vision (ECCV)*. 2024.

[5] J. L. Schönberger and J.-M. Frahm. “Structure-from-Motion Revisited”. In: *Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.

[6] S. Shekhar, M. Reimann, J. I. Wattaseril, A. Semmo, J. Döllner, and M. Trapp. “ALIVE: Adaptive-Chromaticity for Interactive Low-light Image and Video Enhancement”. In: (2023).

[7] Z. Zhang, W. Sun, X. Min, T. Wang, W. Lu, and G. Zhai. “No-Reference Quality Assessment for 3D Colored Point Cloud and Mesh Models”. In: *IEEE Transactions on Circuits and Systems for Video Technology* (2022), pages 1–1. DOI: `10.1109/TCSVT.2022.3186894`.