Internship Report

HUANG Julien

Department **VRU**
10/06/2024 - 12/08/2024

1 Acknowledgments

I would like to express my sincere gratitude to the following individuals and groups who have contributed to the success of my research internship:

- Professor Liu, who assigned me to a team, led our meetings, and provided direction for my tasks.
- Wujiabo, my internship mentor, who was always there to assist me whenever I had questions or needed guidance.
- Haotong, who warmly welcomed me to the university, helped me with various administrative procedures, and even picked me up from the airport.
- The VRU team, who welcomed me into their group and with whom I had the pleasure of attending meetings.
- Vivian Yan, for handling all the administrative processes prior to my arrival at HUST.

Contents

List of Figures

2 Introduction

2.1 Objectives

The primary objective of the team is to develop a spacetime version of Scaffold-GS with a significant reduction in the data size required for video processing, aiming for efficient compression. Their main focus is on minimizing the data footprint while maintaining high-quality results.

Weekly meetings are held to discuss progress and challenges, conducted primarily in Chinese.

Another important goal is to create a viewer application for their newly developed method, allowing for effective visualization and evaluation of the results.

Figure 1: HUST

3 Literature Review

3.1 Overview of Related Work

3.1.1 NeRF : *Neural radience field*

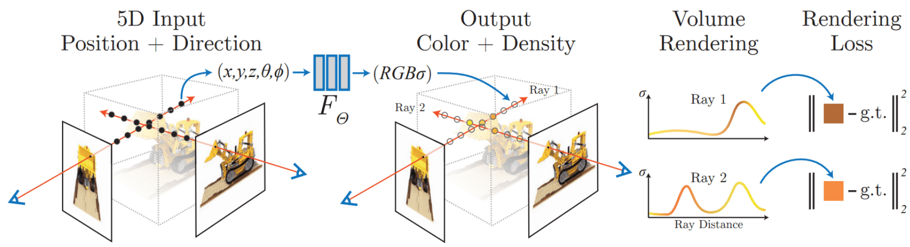

Figure 2: NeRF

NeRF, or Neural Radiance Field, is a method for synthesizing novel views of complex 3D scenes based on a sparse set of 2D images.

3.1.2 3DGS : *3D Gaussian Splatting*

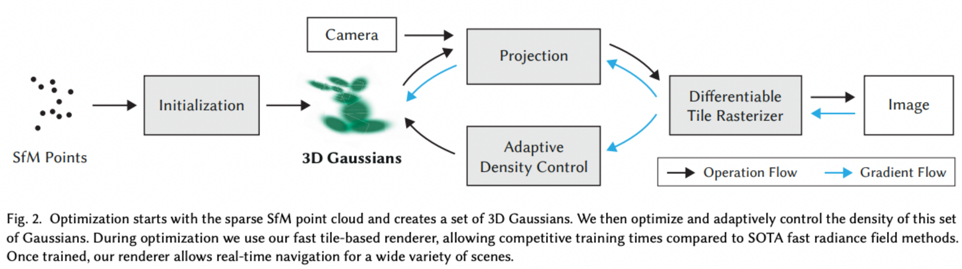

Figure 3: 3DGS

3D Gaussian Splatting (3DGS) is a technique that represents a scene using a collection of 3D Gaussians, each associated with a specific position, color, and opacity.

3.2 Current Trends

3.2.1 Spacetime *Gaussian Feature Splatting for Real-Time Dynamic View Synthesis*

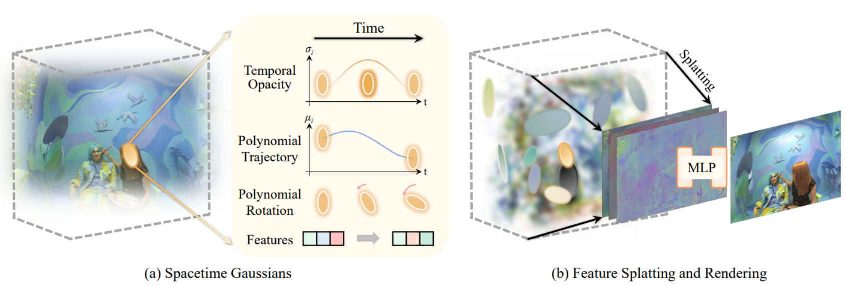

Figure 4: Spacetime

Spacetime Gaussian Feature Splatting is an emerging technique designed for real-time dynamic view synthesis. It extends the 3D Gaussian Splatting approach by incorporating temporal information, allowing for the rendering of dynamic scenes where objects and camera positions change over time.

3.2.2 Scaffold-GS : *Structured 3D Gaussians for View-Adaptive Rendering*

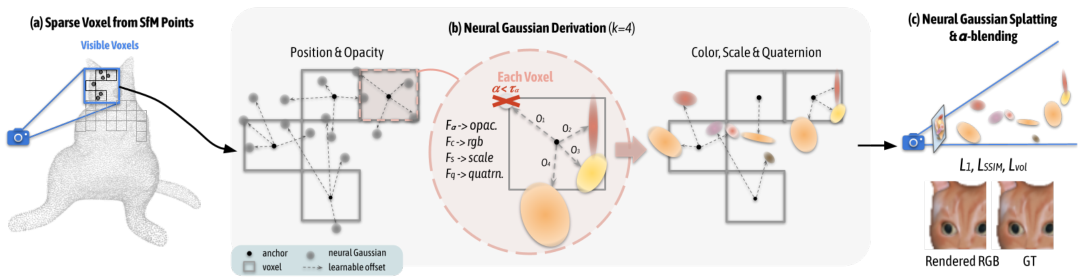

Figure 5: Scaffold-GS

Scaffold-GS introduces a structured approach to 3D Gaussian splatting, focusing on view-adaptive rendering. By organizing 3D Gaussians into a structured grid.

3.2.3 HAC : *Hash-grid Assisted Context for 3D Gaussian Splatting Compression*

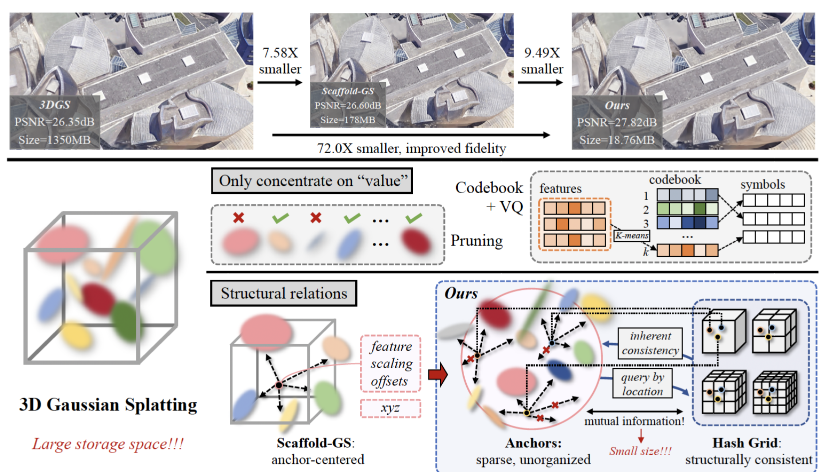

Figure 6: HAC

The Hash-grid Assisted Context (HAC) technique focuses on the compression of 3D Gaussian splatting data. By leveraging a hash-grid structure, HAC effectively reduces the storage requirements for 3D Gaussian representations without compromising on rendering quality.

3.3 Gaps and Challenges

One of the primary challenges I faced was reading and comprehending research papers for the first time. This process involved understanding complex concepts and methodologies that were often presented in dense, technical language. Additionally, conducting in-depth research on these advanced topics required a significant amount of time and effort to grasp the underlying principles and state-of-the-art techniques.

4 Methodology

4.1 Research Design

All the necessary information and resources for this research are available on GitHub, where the project repository is maintained.

4.2 Data Collection

For the static version (without the temporal dimension), a series of photos were taken from different viewpoints. To incorporate the fourth dimension (time), videos were captured from various perspectives. These datasets are essential for constructing a Structure from Motion (SfM) model, which generates a point cloud with various properties.

4.3 Implementation

The team modified the data structure of the Gaussian cloud, selectively storing the relevant information, such as the positions of anchor points, color, rotation, motion, and Multi-Layer Perceptron (MLP) parameters.

5 Weekly Progress Reports

5.1 Week 1

5.1.1 Objectives

The primary objective for the first week was to familiarize myself with the university environment and adjust to the new lifestyle.

Figure 7: Spicy food from Wuhan

5.1.2 Activities

During this week, I focused on settling into the dormitory. I also acquired a Chinese SIM card and unlocked access to the campus's electric scooters, which are essential for getting around, given the large size of the campus. Traveling to the lab on foot takes about an hour, but only 10 minutes by scooter.

5.1.3 Results and Challenges

One of the challenges I faced was learning how to ride an electric scooter, especially since I don't have a driver's license or prior knowledge of traffic rules. Additionally, I familiarized myself with essential apps for daily life in China, such as WeChat and Alipay.

Figure 8: Dorm

Figure 9: Student's ID

Figure 10: Scooter

5.2 Week 2

5.2.1 Objectives

The main objective for the second week was to read and understand various scientific papers to gain a deeper insight into the research objectives and challenges.

5.2.2 Activities

I focused on studying key topics such as NeRF, Gaussian Splatting, Scaffold-GS, and Spacetime.

5.2.3 Results and Challenges

By watching numerous YouTube videos, I was able to grasp the key fundamentals of these research papers, which significantly helped in understanding the complex concepts involved.

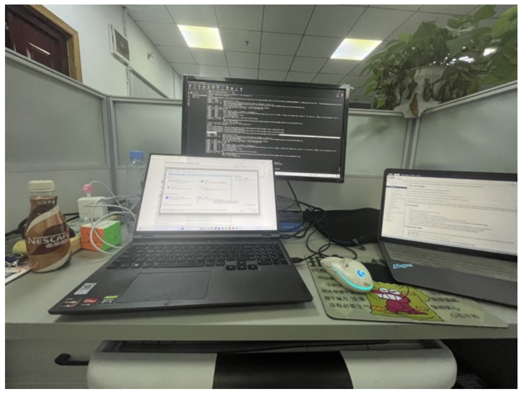

Figure 11: My desk

5.3 Week 3

5.3.1 Objectives

The objective for the third week was to familiarize myself with the SIBR player by building it from the source files. This would allow me to visualize static scenes using Gaussian splatting and make modifications to incorporate the temporal dimension for visualizing Spacetime GS.

5.3.2 Activities

I worked with new tools such as SSH, MobaXterm, and Visual Studio 2019 to accomplish the tasks.

5.3.3 Results and Challenges

One of the main challenges was installing all the necessary resources, which required very specific versions to function correctly.

5.4 Week 4

5.4.1 Objectives

The objective for the fourth week was to analyze the SIBR source code to deduce a pipeline. This would enable us to later modify specific sections to display the lab's research results.

5.4.2 Activities

I began learning a new programming language, C++, primarily through YouTube videos.

5.4.3 Results and Challenges

Since the source code was written in C++, this presented a significant challenge. Although a week is not enough time to fully master a new programming language, I believe it was sufficient to grasp the key concepts within the source code.

5.5 Week 5

5.5.1 Objectives

The main objective for the fifth week was to deduce a pipeline from the SIBR source code.

5.5.2 Activities

To streamline my work, I installed new tools. Sourcetrail helped me navigate the files more effectively and provided better analysis of the dependencies between them. Additionally, I used Beyond Compare 5 to simplify the comparison of source codes between different players, including Gaussian Splatting (original), Scaffold-GS, and Spacetime, in order to identify the key methods that need modification.

I also prepared and delivered a presentation, which was conducted in Chinese.

Figure 12: Meeting

https://github.com/juleelee/internship-hust

5.6 Week 6

5.6.1 Objectives

The objective for the sixth week was to get familiar with the available Spacetime viewer online (https://antimatter15.com/splaTV/).

5.6.2 Activities

I attempted to create a 3D model of myself. However, the result was not very satisfying, which was expected since I only completed 7,000 iterations instead of the recommended 30,000 due to the excessive time it would have taken.

5.6.3 Results and Challenges

I continued to face challenges during the process. Working over SSH complicates the task of downloading the same versions of tools and software. It would have been simpler if I had access to a Windows machine.

5.7 Week 7

5.7.1 Objectives

The objective for the seventh week was to analyze the available online player for 3D Gaussian splatting (https://juleelee.github.io/splat/).

5.7.2 Activities

I analyzed the Python script that converts .ply files (our Gaussian point clouds) into .splat files, a format specifically designed for Gaussian splatting. I then attempted to adapt this script to create a new one that could convert .ply files into .splav files, which are tailored for the Spacetime method (https://antimatter15.com/splaTV/).

5.7.3 Results and Challenges

I concluded that this task is too complex. Developing a new player that combines Scaffold-GS and Spacetime methods is particularly challenging. The Scaffold method uses MLPs (Multi-Layer Perceptrons) to determine various characteristics of the Gaussians over time, such as position, color, rotation, and opacity. While players exist for the Spacetime method, there is no player available for the Scaffold method.

My initial idea was to draw inspiration from existing players (written in JavaScript) and merge their functionalities. However, given that no player currently exists for the Scaffold method and considering the complexity of the task, this mission is beyond my capabilities as an individual. It's worth noting that the Spacetime method player was implemented by someone at MIT, highlighting the level of expertise required.

5.8 Week 8

5.8.1 Objectives

I set a personal objective to create a website that allows users to compare images with the ground truth, accessible from any computer.

5.8.2 Activities

I worked on calculating PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index) to quantify the differences between the images and the ground truth. I also implemented thresholds to highlight these differences.

5.8.3 Results and Challenges

From my perspective, the result was a satisfactory website. The main challenges were creating a site that simplifies the process of visually comparing the results of different methods, and ensuring that the site was user-friendly and easy to navigate.

Figure 13: Compare

Figure 14: Compare

6 Conclusion

6.1 Summary

During this internship, I explored the advanced concepts of 3D Gaussian Splatting, including NeRF, Scaffold-GS, and Spacetime. Through reading scientific papers, analyzing source code, and hands-on experimentation, I gained valuable insights into 3D rendering and dynamic view synthesis.

6.2 Achievements

Key achievements include adapting the SIBR player and developing a website for comparing rendered images with ground truth. Additionally, I built a foundational understanding of C++, which was essential for working with the SIBR source code. These accomplishments have greatly enhanced my technical skills in 3D graphics research.

6.3 Team Cohesion and Cultural Experiences

Beyond the technical work, I also had the opportunity to engage in team-building activities and explore the city of Wuhan. These experiences were invaluable for fostering team cohesion and allowed me to gain a deeper appreciation of the local culture. Below are some images capturing moments from these outings and team activities.

Figure 15: Hotpot with the team

Figure 16: Mahjong

Figure 17: Yangzi river

Figure 18: City

Figure 19: Yellow Crane Tower