

Simplifying NeRF: Creating an Intuitive Web-Based 3D Scene Interface

Eduard von Briesen

March 31, 2023
Version: My First Draft



Department of Mathematics,
Informatics and Statistics,
Institute of Informatics



Munich Film School,
Chair for AI

Masters Thesis

Simplifying NeRF: Creating an Intuitive Web-Based 3D Scene Interface

Eduard von Briesen

Supervisors

Prof. Dr. Sylvia Rothe and Cristoph Weber

March 31, 2023



Eduard von Briesen

Simplifying NeRF: Creating an Intuitive Web-Based 3D Scene Interface

Masters Thesis, March 31, 2023

Supervisors: Prof. Dr. Sylvia Rothe and Cristoph Weber

LMU Munich

Department of Mathematics, Informatics and Statistics

Institute of Informatics

Artificial Intelligence and Machine Learning (AIML)

Akademiestraße 7

80799 Munich

Abstract

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Abstract (German)

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Acknowledgement

Contents

- 1. Introduction 1**
 - 1.1. Background 1
 - 1.2. Overview of NeRF 2
 - 1.3. Research Objectives 2
 - 1.4. Research Question 3
- 2. Related Work 5**
 - 2.1. Existing Methods for creating NeRFs 5
 - 2.2. Review of tools and technologies in film and VFX 5
 - 2.3. User Research 5
 - 2.4. Conclusion 5
- 3. Methodology 7**
 - 3.1. Research Design 7
 - 3.2. Initial User Research 7
 - 3.2.1. Participant Selection Criteria 7
 - 3.2.2. Interview Methodology 8
 - 3.2.3. Key Findings 8
 - 3.3. Prototype Development 10
 - 3.3.1. Technical Specifications 10
 - 3.3.2. Design Considerations 10
 - 3.4. User Study and Evaluation 10
 - 3.4.1. Tasks Based Usability Test 10
 - 3.4.2. Follow-up Interview 11
 - 3.4.3. Qualitative Analysis 11
- 4. Technical Implementation 13**
 - 4.1. System Architecture 13
 - 4.2. Frontend Development 13
 - 4.3. Backend Development 14
 - 4.4. Challenges and Solutions 14
 - 4.5. Lessons Learned 14
 - 4.6. Future Directions 14

5. Results	15
5.1. Analysis of User Experience Questionnaire	15
5.2. Findings from Qualitative User Testing	15
5.3. Integration and Findings	15
6. Conclusion	17
6.1. Key Findings	17
6.2. Contributions to the Field	17
6.3. Future Work	17
A. Appendix	19
A.1. Interview Questions	19
A.2. User Study Questionnaire	21
A.3. User Testing Results	22
A.4. Prototype Documentation	22
Bibliography	23
List of Figures	25
List of Tables	27

Introduction

1.1 Background

The present state of NeRF research presents significant advancements that have greatly influenced the field of 3D scene modeling and rendering. This section provides an overview of the current research landscape, highlighting key NeRF frameworks and relevant projects, and identifying the challenges and opportunities that inform our research objectives.

Two prominent NeRF frameworks with user interfaces, namely Instant NGP [Mül+22] and Nerfstudio [Tan+23], have emerged as leaders in enabling users to explore and manipulate 3D scenes. These frameworks offer features such as real-time scene rendering, adjustable training parameters, and the creation of camera trajectories for video rendering.

However, the utilization of these interfaces often demands a high level of technical expertise, as they are designed to complement, rather than replace, command-line interfaces. Users must engage with terminal-based processes for tasks such as video data preprocessing, model training, and rendering output.

Additionally, several innovative projects have expanded the NeRF landscape. Notably, CLIP-NeRF [Wan+22], Instruct-NeRF2NeRF [Haq+23], Text2LIVE [Bar+22], and SINE [Bao+23] have introduced text-based editing approaches, broadening the possibilities for manipulating NeRF models. PaletteNeRF [WTX22] focuses on color editing, while NeRF-Editing [Yua+22] enables mesh editing.

This research aims to identify key challenges and opportunities in NeRF frameworks and interfaces, as demonstrated by these significant contributions. This knowledge will guide the development of a user-friendly, web-based interface and integrated editing plugins, with the ultimate goal of enhancing the accessibility and usability of NeRF frameworks for a broader user base.

1.2 Overview of NeRF

1.3 Research Objectives

The research objectives of this study are as follows:

1. **Exploration of NeRF Interaction Capabilities:** This study aims to explore the existing interaction capabilities within NeRF frameworks comprehensively. It involves an analysis of the current state of NeRF interfaces and an investigation into user engagement, visualizations, and manipulation of NeRF scenes.
2. **Development of a Web-Based User Interface:** Building on insights gained from the exploration phase, the primary objective is to design and implement a user-friendly web-based interface for NeRF.
3. **Streamlined NeRF Creation and Manipulation:** The central goal is to simplify the process of NeRF creation and manipulation, eliminating the need for users to deal with complex command-line interfaces or extensive local setup. The web-based interface will provide an intuitive and efficient user experience.
4. **Integration of Diverse Editing Plugins:** To enhance the creative potential of NeRF, various editing plugins will be integrated into the web-based interface. The objective is to expand the functionality and versatility of the NeRF framework.

The research aims to advance NeRF frameworks' capabilities and accessibility, making them accessible to a broader audience and fostering innovation in 3D scene modeling and rendering.

1.4 Research Question

1. **Enhancing NeRF Frameworks:** How can a web-based interface improve the user experience and accessibility of NeRF frameworks, and what impact will these enhancements have on user-friendly NeRF creation and manipulation?
2. **Overcoming Technical Challenges:** What technical challenges and limitations are associated with current NeRF frameworks and interfaces, and how can innovative design and technology choices in a web-based interface overcome these challenges?
3. **Innovative Editing Integration:** How can novel editing approaches be seamlessly integrated into a web-based NeRF interface to enhance creativity and usability, and how do these methods compare with traditional NeRF editing techniques?

Related Work

2.1 Existing Methods for creating NeRFs

Luma AI

Nerfstudio

2.2 Review of tools and technologies in film and VFX

2.3 User Research

2.4 Conclusion

Methodology

3.1 Research Design

The research was planned in three distinct phases, initial user research, development of the prototype and user testing. Each of these phases informs the one building on it.

3.2 Initial User Research

The initial round of user interviews was conducted with to gather insights for the development of a user-friendly NeRF interface. The focus was on understanding the needs, challenges and preferences of users with varying levels of experience with the creation of NeRF models.

3.2.1 Participant Selection Criteria

Participant were selected based on there pre-existing experience with NeRF and some association to the film industry. They represent a wide range of experiences and technical knowledge of NeRF, but also applied the technology in different contexts.

This diversity in users helped to identify common needs and challenges, as well as to understand the specific requirements of different user groups.

3.2.2 Interview Methodology

The interviews themselves were semi-structured, following a set of predefined questions (A.1), while also allowing for open-ended discussions and follow-up questions. The interviews were conducted in a one-on-one setting, either in person or via video call, and lasted between 30 and 60 minutes. Although the interview template was prepared in english, all interviews were conducted in german.

Questions followed a logical flow, starting with general background information and moving on to more specific needs and preferences. After being introduced to the purposes of the interview and the goals of this research, participants were questioned about their experience with NeRF and the tasks they typically aim to achieve when working with NeRF. They were asked what challenges or pain points they encounter when using current NeRF frameworks or interfaces, and what features or functionalities would make a NeRF interface most useful for their work or research. Interviewees were encouraged to share there own thoughts and suggestions on the development of a NeRF interface.

With the agreement of participants, the interviews were recorded and transcribed for further analysis. The data was then coded and categorized to identify common themes and patterns across the interviews.

3.2.3 Key Findings

NeRF in the Film Industry

There exist different applications of NeRF in the film industry, ranging from traditional visual effects to virtual production and location scouting in the pre-production phase. Tradition creation and animation of 3D scenes is costly and time-consuming, and NeRF is seen as a potential solution to this problem.

Some participants expressed their concerns about the quality of current NeRF models, that is often not sufficient for professional use. Lack of detail and visual artifacts in the scene were mentioned as common issues. Additionally it lacks the ability to edit and manipulate the models when compared to traditional 3D models,

which presents a crucial limitation for the application in a creative industry like film, where models are iteratively developed and refined.

A more realistic use-case of NeRF in the film industry, with the technologies current limitations, is the use in pre-visualization and location scouting. The ability to quickly capture and render 3D scenes from real-world locations is seen as a potential time-saver in the pre-production phase of film-making. Models can be used to plan camera movements and lighting, and to get a better understanding of the spatial relations between different elements in the scene. One concern raised in this context was the lack of accurate scale of the models, when exporting them into other 3D software.

Optimizing Parameters and Workflow

A typical workflow when creating NeRFs follows three steps, first the input data has to be pre-processed, then the actual model training is done and finally the some data is exported. Parameterization in the pre-processing and training phases can have a profound impact on the quality of the final NeRF. This was noted by participants with a technical background, that were working with NeRF in a research context. For them the optimization of parameters is an integral part of their workflow, often re-training models multiple times, in order to analyze the variation in the results. Results can be quantified with tooling to analyze runs, such as TensorBoards.

User Interface and User Experience

Participants

3.3 Prototype Development

3.3.1 Technical Specifications

3.3.2 Design Considerations

3.4 User Study and Evaluation

To evaluate the usability and usefulness of the developed prototype, a user study was conducted. The study was designed to gather feedback on the user experience of the prototype, and to identify any usability issues or challenges that users encountered. The study consisted of task based usability tests, followed by a short follow-up interview.

3.4.1 Tasks Based Usability Test

The usability test was conducted in a controlled environment, with participants being asked to complete a two tasks with the prototype. The tasks were designed to cover a range of functionalities and features of the prototype, and represent a typical workflow when creating NeRF models.

1. Task: Create a NeRF model from a set of input images
2. Task: Export the created NeRF model into a video.
3. Task: Edit the created NeRF model using prompts.

To keep an appropriate time frame, none of the tasks required completion of a training process, and pre-processed data and pre-trained models were provided to the participants. On average, participants took 30 minutes to complete the tasks.

Participant were passively observed while working on their tasks, to identify any problems or operation errors they encountered and to determine their overall per-

formance. In addition the screen was recorded to capture the participants interactions with the prototype, and to allow for a more detailed analysis of their behavior later on.

3.4.2 Follow-up Interview

After completing the usability test, participants were engaged in a short follow-up interview, to gather more detailed feedback on their experience with the prototype. Similar to the initial user interviews, these interviews were semi-structured, following a pre-defined set of question, with room for participants to share their own thoughts and suggestions. The questions can be categorized into general usability, tasks specific feedback and suggestions for improvement. The interview template can be found in the appendix A.2.

3.4.3 Qualitative Analysis

Technical Implementation

4.1 System Architecture

The architecture followed a standard client server pattern, with the server functioning as a wrapper for the nerfstudio CLI, and the client as a web application. The server was built using tRPC, a framework for building type-safe APIs in TypeScript. It was responsible for handling incoming requests from the client, and translating them into commands that the nerfstudio CLI could understand. Requests from the client were sent to the server using HTTP requests, and in case of an asynchronous operation, the server would update the client on using websockets.

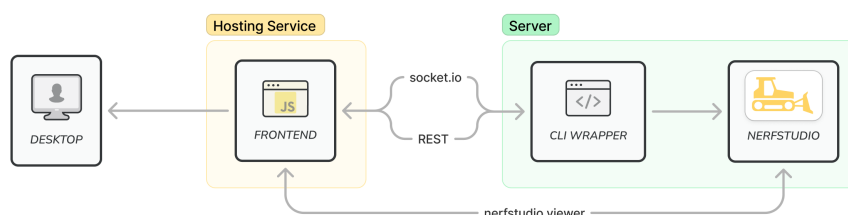


Fig. 4.1.: System Architecture

For the prototype, all the components were composed into a single docker container and deployed as a single unit. This leveraged the pre-configured container provided by nerfstudio, and allowed for a quick and easy deployment of the prototype.

4.2 Frontend Development

The frontend was built with React, as it is a popular and well supported framework for building web applications. Vite was used as a build tool, as it provides a fast

and efficient development experience. To accelerate the styling process, Tailwind CSS was used, it is a utility-first CSS framework that provides a set of pre-defined classes that can be used to style components. In addition the daisyUI component library provided a set of pre-styled components that could be used to quickly build the UI.

Extensibility was a key consideration during the development of the frontend, ...

4.3 Backend Development

The backend was built using tRPC, a framework for building type-safe APIs in TypeScript. This type-safety was useful in building the API, as nerfstudio endpoints require a specific set of parameters of various types, that could be easily defined using TypeScript and reused in the frontend.

With tRPC,

4.4 Challenges and Solutions

4.5 Lessons Learned

4.6 Future Directions

Results

5.1 Analysis of User Experience Questionnaire

5.2 Findings from Qualitative User Testing

5.3 Integration and Findings

Conclusion

6.1 Key Findings

6.2 Contributions to the Field

6.3 Future Work

Appendix

A

A.1 Interview Questions

Introduction

- Thank the interviewee for their participation.
- Briefly explain the purpose of the interview, which is to gather insights for the development of a NeRF interface.
- Assure the interviewee that their responses will be kept confidential.

Background

- Can you briefly describe your experience with NeRF or 3D modeling in general?

Needs and Challenges

- What specific tasks or objectives do you typically aim to achieve when working with NeRF models or 3D scenes?
- What are the main challenges or pain points you encounter when using current NeRF frameworks or interfaces?

Usability and Features

- In your opinion, what features or functionalities would make a NeRF interface most useful for your work or research?
- Are there any specific editing or manipulation tools you find lacking in current NeRF interfaces?

- How important is real-time visualization and interactivity in a NeRF interface for your needs?

Ease of Use

- How do you envision the ideal user interface for NeRF? What elements would make it easy to use, even for those with limited technical expertise?
- What level of technical knowledge or familiarity with 3D modeling should the ideal NeRF interface require from its users?

Integration and Compatibility

- Are there any other software tools or workflows you typically use alongside NeRF, and how important is it for a NeRF interface to integrate with these tools?
- Do you have any preferences regarding the file formats or data compatibility that the NeRF interface should support?

Feedback and Suggestions

- Are there any additional thoughts, suggestions, or requirements you would like to share regarding the development of a NeRF interface?
- Is there anything else you believe is essential for us to understand about your needs and expectations?

Closing

- Thank the interviewee for their time and valuable input.
- Provide contact information for any follow-up questions or clarifications.

A.2 User Study Questionnaire

Usability Experience

- Can you share your overall impressions of using the Simplify NeRF application?
- Were there any specific features or functionalities of the application that stood out to you positively? If so, why?
- On the other hand, were there any aspects of the application that you found challenging or frustrating? Please elaborate.

Task-Specific Feedback

- Were there any particular steps or actions within the tasks that you found confusing or unclear? If yes, could you describe them?
- Were there any features or functionalities you expected to find in the application that were missing? If so, what were they?

Suggestions for Improvement

- Based on your experience using the Simplify NeRF application, do you have any suggestions for improving its usability or functionality?
- Are there any specific changes or enhancements you would like to see in future versions of the application?
- How do you think the application could better meet your needs or expectations as a user?

Closing

- Is there anything else you would like to add or share about your experience with the Simplify NeRF application?

- Thank the participant for their time and valuable feedback. Offer contact information for any follow-up questions or clarifications.

A.3 User Testing Results

A.4 Prototype Documentation

Bibliography

- [Bao+23] Chong Bao, Yinda Zhang, Bangbang Yang, et al. “SINE: Semantic-driven Image-based NeRF Editing with Prior-guided Editing Field”. In: *The IEEE/CVF Computer Vision and Pattern Recognition Conference (CVPR)*. 2023 (cit. on p. 1).
- [Bar+22] Omer Bar-Tal, Dolev Ofri-Amar, Rafail Fridman, Yoni Kasten, and Tali Dekel. *Text2LIVE: Text-Driven Layered Image and Video Editing*. 2022. arXiv: 2204.02491 [cs.CV] (cit. on p. 1).
- [Haq+23] Ayaan Haque, Matthew Tancik, Alexei Efros, Aleksander Holynski, and Angjoo Kanazawa. “Instruct-NeRF2NeRF: Editing 3D Scenes with Instructions”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023 (cit. on p. 1).
- [Mül+22] Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. “Instant Neural Graphics Primitives with a Multiresolution Hash Encoding”. In: *ACM Trans. Graph.* 41.4 (July 2022), 102:1–102:15 (cit. on p. 1).
- [Tan+23] Matthew Tancik, Ethan Weber, Evonne Ng, et al. “Nerfstudio: A Modular Framework for Neural Radiance Field Development”. In: *ACM SIGGRAPH 2023 Conference Proceedings*. SIGGRAPH ’23. 2023 (cit. on p. 1).
- [Wan+22] Can Wang, Menglei Chai, Mingming He, Dongdong Chen, and Jing Liao. “CLIP-NeRF: Text-and-Image Driven Manipulation of Neural Radiance Fields”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2022, pp. 3835–3844 (cit. on p. 1).
- [WTX22] Qiling Wu, Jianchao Tan, and Kun Xu. *PaletteNeRF: Palette-based Color Editing for NeRFs*. 2022. arXiv: 2212.12871 [cs.CV] (cit. on p. 1).
- [Yua+22] Yu-Jie Yuan, Yang-Tian Sun, Yu-Kun Lai, et al. “NeRF-Editing: Geometry Editing of Neural Radiance Fields”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2022, pp. 18353–18364 (cit. on p. 1).

List of Figures

4.1. System Architecture 13

List of Tables

Colophon

This thesis was typeset with \LaTeX 2_ε. It uses the *Clean Thesis* style developed by Ricardo Langner. The design of the *Clean Thesis* style is inspired by user guide documents from Apple Inc.

Download the *Clean Thesis* style at <http://cleanthesis.der-ric.de/>.

