Concept-Infused 3D Editing of Neural Radiance Fields

Jinang Shah

Stanford University

Motivation

The project, "Concept-Infused 3D Editing of Neural Radiance Fields," addresses a critical need for enhanced 3D editing capabilities with real-world applications in virtual reality, augmented reality, and computer graphics. Current limitations in 3D editing tools restrict detailed customization, impacting industries like gaming and architecture. By infusing specific visual concepts into the editing process, the project aims to empower users with advanced control, fostering creativity and improving immersive experiences. The practical implications of solving this problem extend to improved content creation, realistic virtual environments, and a more engaging user experience across various domains.

Related Work

View-Consistency-Aware 3D Editing of Neural Radiance Fields (ViCA-NeRF) [2] is the first view-consistency-aware method for 3D editing that utilizes recent advancements in text-to-image models [3] and diffusion models [5] to perform text instructions guided 3D editing, producing notable results. However, the results of the 3D edits are not customizable beyond the text prompts and tend to omit high-level details. Concept Decomposition [1] provides tree-structured textual vector embeddings that enable capturing specific visual concepts, whereas [4] learns unique concepts and style through converting given images into its word embeddings. But this technique has not yet been applied to 3D customization and editing as a guiding process. To increase the quality of 3D reconstructions and the degree of control, we intend to refine view-consistency-aware 3D editing of NeRFs by adding the ability to “infuse” visual concepts in NeRF editing.

Approach

To enhance concept-infused 3D editing of NeRFs, we propose a dual approach. First, we generate concept images or embeddings from the original images using the concept decomposition model [1]. These representations capture specific visual concepts, bridging the gap between 2D images and 3D editing [2]. For instance, given an original image and a concept (e.g., "Iron Man"), we create concept images of person wearing Iron Man suit, Iron Man, suit, etc., and feed them into ViCA-NeRF framework for reference. Second, as an alternative, we directly integrate the textual vector embeddings from the concept decomposition model into the ViCA-NeRF framework for 3D editing. We adapt the Instruct-Pix2Pix model to accept the concept images or embeddings, ensuring multi-view consistency through geometric and learned

* This project incorporates insights from collaborative work with Jinhyo Huh, a fellow master's student, contributing to a broader research initiative.

regularization strategies. Additionally, we employ warm-up, blending, and post-refinement techniques to enhance editing quality and efficiency.

Expected Results

For the approach of first concept-infusing the 2D images and then using them to 3D edit, we expect the final 3D scene to have a more consistent level of detail that strongly adheres to the original concept image. This would allow fine customization of character and object aspects in a 3D context, not just in 2D. For the approach of taking the text embeddings directly from the to produce the text-generated 3D edits, the results may be more prone to turbulence and artifacts and may not differ too much from the initial paper. For both approaches, we expect the speed of the compute to be slower than the current ViCA-NeRF, as we are adding the steps of concept decomposition, then generating concept-infused images and 3D edits.

References

[1] Concept Decomposition for Visual Exploration and Inspiration

[2] ViCA-NeRF: View-Consistency-Aware 3D Editing of Neural Radiance Fields

[3] Learning Transferable Visual Models From Natural Language Supervision

[4] An Image is Worth One Word: Personalizing Text-to-Image Generation using Textual Inversion

[5] High-resolution image synthesis with latent diffusion models

* This project incorporates insights from collaborative work with Jinhyo Huh, a fellow master's student, contributing to a broader research initiative.