Toon NeRF

🕐 Created	@November 16, 2022 3:03 PM
🕐 Last Edited Time	@November 16, 2022 4:52 PM
⊙ Type	
⊙ Created By	🖼 Joonho Kim

Motivation

The animation industry has traditionally spent a majority of time animating novel views of characters manually using hand-drawn techniques. More recently with Computer Graphics techniques, geometric representations such as meshes paired with rendering algorithms can automate novel views of characters and scenery. However, much of these geometric representations take time to construct and lose the fidelity of cartoon exageration where as hand-drawn images maintain an artists orginal vision of a character. Is there ways that modern novel view synthesis learning techniques could be applied into the artistic domain? Specifically, could you apply NeRF's to cartoon character drawings and generate novel views of those characters?

Related Work

Novel view synthesis tackles the problem of rendering any desired view of a subject given only a subset of views and respective renders. Recently the volumetric-based representation Neural Radiance Fields [2] has gained much popularity by representing a subject as a density field and view-dependent radiance. Much followup work has driven NeRF's to handle deformations [4], faster representations [5], and fewer input [6].

Novel view generation of cartoon images within the 2D graphics community have taken more geometric approaches such as with rotating planar representations [1] or image feature matching/interpolation [7]. The most widely sought after application of novel view generation is inbetweening (drawing frames inbetween two key frames), which could use the application of deformable NeRfs given few character inputs and real-time

manipulation; however this idea has not been published or explored in depth with NeRFs.

Project Overview

I have been looking at the idea of cartoon character inputs for NeRF's for the past couple of months but more recently I have began experimenting with implementations and ideas. The main issues of this problem are the following:

- Few Shot Input: The artist can realistically provide very few input images for the NeRF to train on. Regular NeRF's don't handle few shot well.

- Cartoon vs Real domains: The cartoon domain is vastly different than the real domain with shading (toon shading) and view dependent geometry (mickey mouse ears effect). Is there ways to capture cartoon characteristics in NeRFs by using cartoon priors.

For my project, I'd like to explore formulating cartoon priors for a NeRF. Because the range of cartoon problems is large, I'd like to specifically focus on generating novel views of a character in a still pose. The raw data can be found in the format of **character reference sheets** where I can estimate camera matrices through overlaying

https://www.artstation.com/artwork/bbZwG
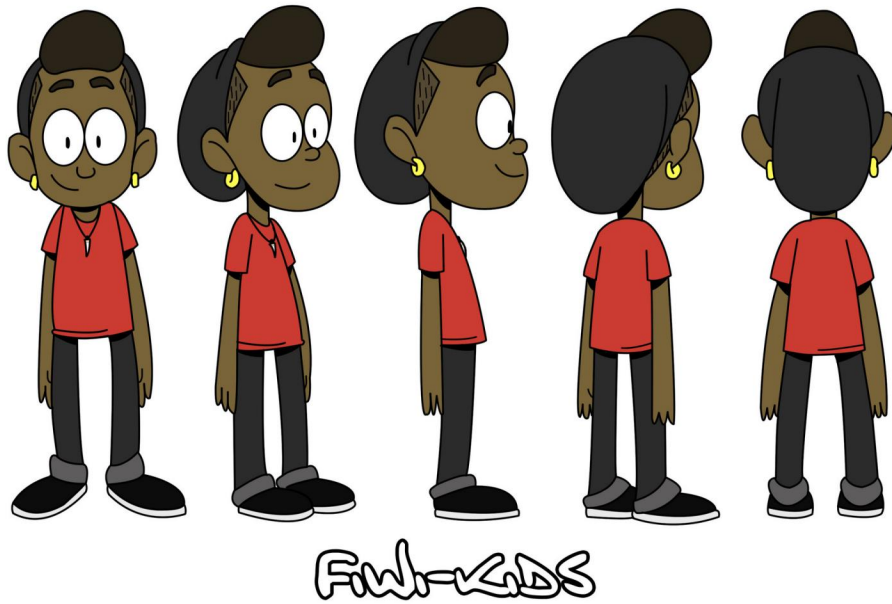
Milestones

My milestones should be very flexible since I am not implementing a previous paper, thus my results are undetermined. But I have a few ideas I can set as milestones for now.

- Implement a color palette regularization with comparisons. Color palettes are commonly used in character design, so the range of rgb values is discrete.

- Image Gradient regularization with comparisons. Cartoon images often have black outlines to indicate segmentation, but implicitly these silhouettes mark surfaces with normals perpendicular to the camera view [3]. We should use this information to regularize the density field of our NeRF.

- Explore backing representations of NeRF's with comparisons: MLP, Hash Tables, Tri-Planar Grids, Voxel Grids.

References

[1] Alec Rivers, Takeo Igarashi, and Frédo Durand. 2010. 2.5D cartoon models. ACM Trans. Graph. 29, 4, Article 59 (July 2010), 7 pages. https://doi.org/10.1145/1778765.1778796

[2] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2021). Nerf: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM, *65*(1), 99-106.

[3] Doug DeCarlo, Adam Finkelstein, Szymon Rusinkiewicz, and Anthony Santella. 2003. Suggestive contours for conveying shape. ACM Trans. Graph. 22, 3 (July 2003), 848–855. https://doi.org/10.1145/882262.882354

[4] Park, K., Sinha, U., Barron, J. T., Bouaziz, S., Goldman, D. B., Seitz, S. M., & Martin-Brualla, R. (2021). Nerfies: Deformable neural radiance fields. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 5865-5874).

[5] Chen, A., Xu, Z., Geiger, A., Yu, J., & Su, H. (2022). TensoRF: Tensorial Radiance Fields. *arXiv preprint arXiv:2203.09517*.

[6] Niemeyer, M., Barron, J. T., Mildenhall, B., Sajjadi, M. S., Geiger, A., & Radwan, N. (2022). Regnerf: Regularizing neural radiance fields for view synthesis from sparse inputs. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5480-5490).

[7] Hojin Choi and Seungkyu Lee. 2017. Novel view synthesis from two cartoon face drawings. In ACM SIGGRAPH 2017 Posters (SIGGRAPH '17). Association for Computing Machinery, New York, NY, USA, Article 8, 1–2. https://doi.org/10.1145/3102163.3102229