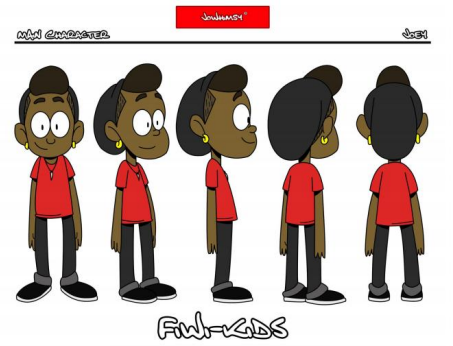
Exploring Cartoon Priors for Neural Radiance Fields

Joonho Kim*
joonho@dgp.toronto.edu, University of Toronto

Motivation

Can we apply cartoon stylization towards learning models to synthesize novel views of a cartoon character?



Joey by JoWayne McFarlane

Animators spend hundreds of hours re-drawing and re-manipulating characters for animation. Often times, the teams will design **character reference sheets** that portray a character's visual attributes. Though humans can understand and visualize new poses from a sparse resource, learning models fail to do the same.

For this project, I explored animation priors and implemented a color-palette prior for NeRFs to try reconstructing characters.

Related Work

2.5D character representations [1] using 2D canvases arranged in 3D. However, this approach only handles simple cartoon models and can not handle rotations of concave silhouettes.

Neural Radiance Fields [2] have become very successful at reconstructing novel views of a scene given a set of images, however NeRFs suffer at few-shot reconstruction for view-dependent geometry, a trait common in exaggerated drawings.

References

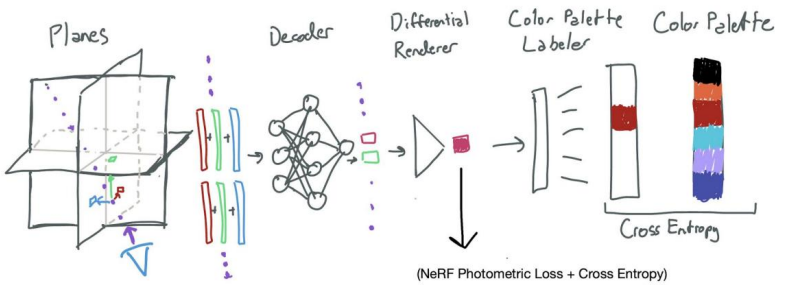
[1] Rivers, Igarashi, and Durand. 2.5D cartoon models. ACM Trans. Graph. 29, 4, Article 59 (July 2010), 2010

[2] Mildenhall, Srinivasan, Tancik, Barron, Ramamoorthi, & Ng, R. Nerf: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM, 2020.

New Technique

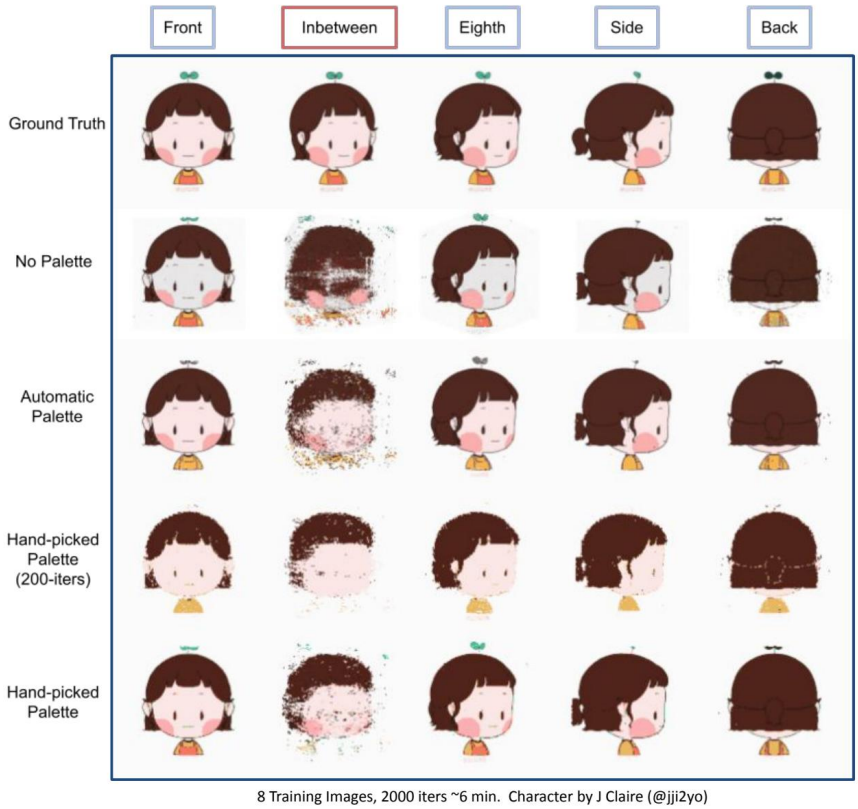
There are many cartoon priors we can explore: flat shading, black silhouettes, and human geometry priors; I applied a discrete color-space prior for this project. Animators often use a restricted color palette when coloring characters, but currently NeRF's predict RGB values from a continuous space.

We first create a color palette manually or from our input images using the Extcolor module. We then classify the RGB pixel values rendered from a NeRF (with planar representation) using a small MLP classifier. Our loss is a sum of NeRF photometric loss and color classification cross entropy loss.



(NeRF Photometric Loss + Cross Entropy)

Experimental Results



8 Training Images, 2000 iters ~6 min. Character by J Claire (@jji2yo)

	PSNR	SSIM	LPIPS
No Palette	22.51	0.778	0.303
No Palette (200 iters)	19.878	0.762	0.35
Palette (automatic)	23.18	0.762	0.246
Palette (hand selected, 200 iters)	17.9	0.767	0.271
Palette (hand selected)	**24.91**	**0.802**	**0.21**

*Project essay will be available on ohnooj.com