3D-Aware anime image synthesis based on π-GAN

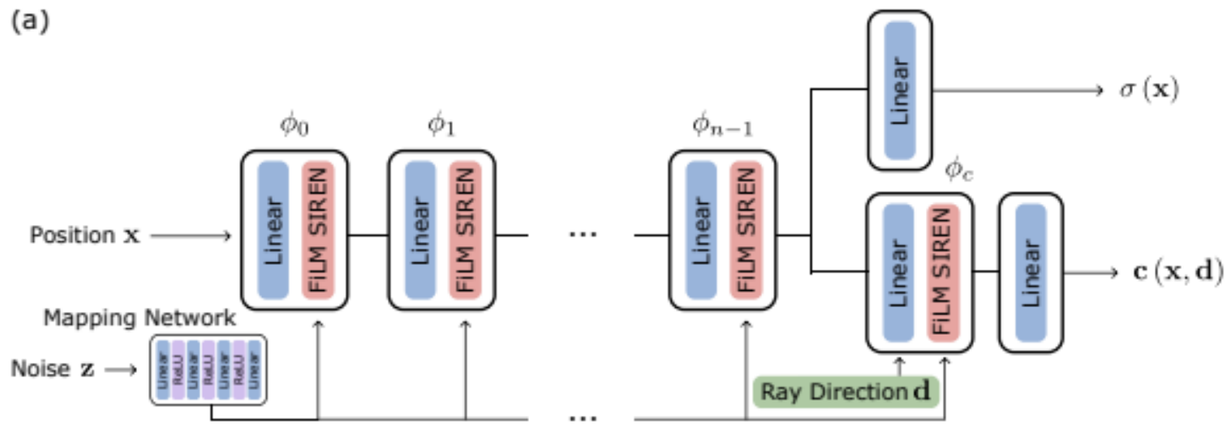
19

曾信彥 陳靖沄 姚舜齊

Introduction

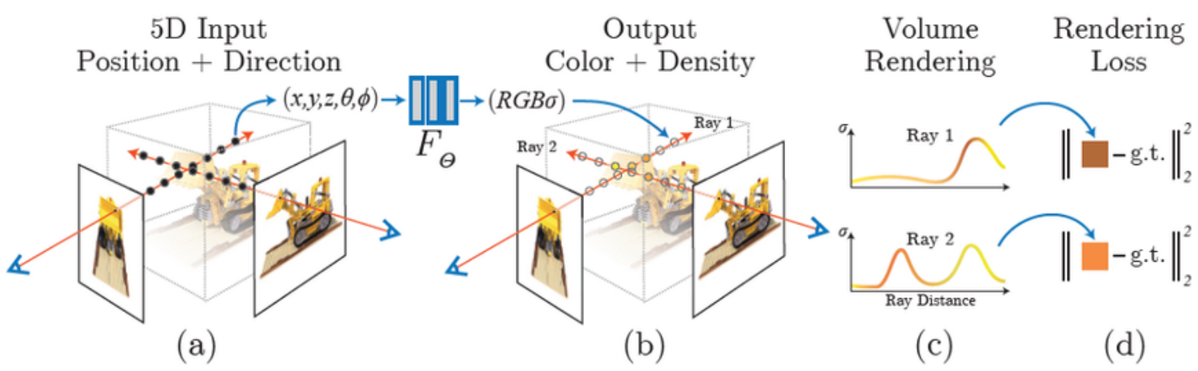
3D-aware images synthesis usually suffers from low resolution or multi-view inconsistency. In order to solve this problem, we introduce a NeRF based model called pi-GAN, which used a Non-voxel based model for 3D representation. This model gets a great result in real-world datasets such as Cats and CelebA, we are wondering if the model works when we apply non-real-world datasets such as Anime dataset.
The dataset consists of 726 animated face images with size 512*512.

Architecture



SIREN-Based Radiance Field and Volume rendering

A radiance field is a continuous mapping from a 3D location and a 2D viewing direction to an RGB color value.



Volume rendering is a technique that render the pixel color C for a camera ray r as follow

$$\mathbf{C}(\mathbf{r}) = \int_{t_n}^{t_f} T(t)\sigma\left(\mathbf{r}(t)\right)\mathbf{c}\left(\mathbf{r}(t),\mathbf{d}\right)dt,$$

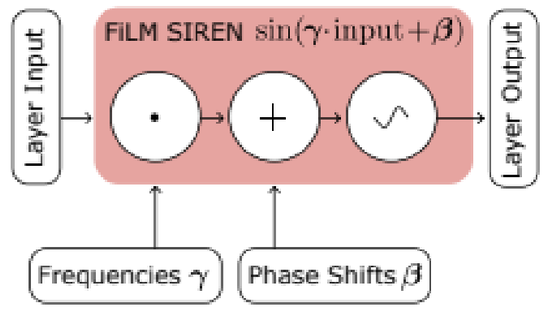
$$\text{where} \quad T(t) = \exp\left(-\int_{t_n}^{t}\sigma(\mathbf{r}(s))ds\right).$$

Our approach implements a discretized form of this equation using the stratified and hierarchical sampling approach introduced by NeRF

$$\mathbf{c}_r = \sum_{i=1}^{N} T_r^i\,\alpha_r^i\,\mathbf{c}_r^i \qquad T_r^i = \prod_{j=1}^{i-1}\left(1-\alpha_r^j\right) \qquad \alpha_r^i = 1-\exp\left(-\sigma_r^i\delta_r^i\right)$$

Training detail

batch size=14
G_lr = 0.00006
D_lr = 0.0002
img_size = 64*64
pretrain model : CelebA



Experiment Result

Discussion

As for the results, we conclude some possible reasons why our 3D anime image cannot be synthesized well.

1. Small dataset with only 737 images.
2. Difference in the image brightness among the dataset may cause black hole on the synthesized image.
3. Among the dataset, some eyes covered by hair are transparent and some are not. It may be the reason why the eyes cannot be synthesized well.
4. Our computational resource is constraint that we can only use batch size=14 instead of 56 in original code. (batch size is important for the beginning of training)

Reference: Eric R. Chan* Marco Monteiro* Petr Kellnhofer Jiajun Wu Gordon Wetzstein Stanford University, "**pi-GAN: Periodic Implicit Generative Adversarial Networks for 3D-Aware Image Synthesis**", CVPR2021