Vector Logo Image Synthesis Using Differentiable Renderer

Ryuta Yamakura†, Keiji Yanai†
†The University of Electro-communications

Introduction

- Recently, parameter optimization techniques are used for vector generation with the differentiable renderer[1].
- It is possible to transfer the knowledge of generative models that handle raster images to vector image generation.
- This study proposes vector format logo image generation using diffusion model knowledge.

Purpose

Generate vector logo images with individual control ;
1. **Outline of the entire image**
2. **Artistic graphics**

Method

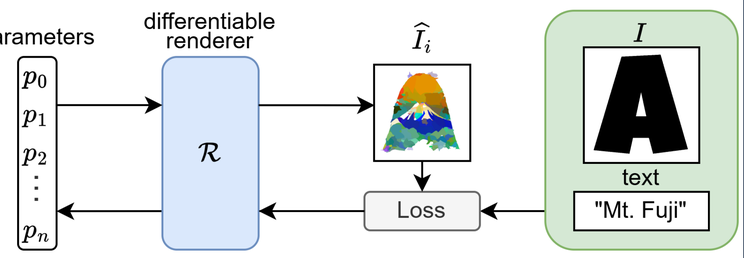
RELATED METHODS

1. Tone Preserving Loss[2] $\mathcal{L}_{tone}$
Limit the region of expression to the range of the input image. MSE between the image after applying a low-pass filter of input and output.

2. Score Distillation Sampling Loss[3] $\mathcal{L}_{LSDS}$
Reflects the content presented in the input text by distilling the knowledge of the diffusion model.

OVERVIEW

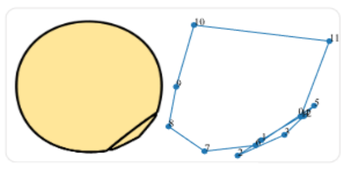
- Represents image using closed Bézier curves.
- Each Bézier curve has parameters[4]; coordinate, color, opacity.
- Optimize Bézier curve parameters using othter condition inputs.



LOSS FUNCTIONS

3. Radiation Loss $\mathcal{L}_{rad}$
Explicitly monitor control stores of Bézier curves.

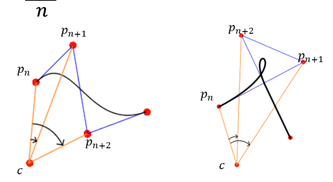
Self-closing problem[5]
The path intersects with itself during the optimization process
Because it only considers pixel spaces



Example of Self-closing (sited from [5])

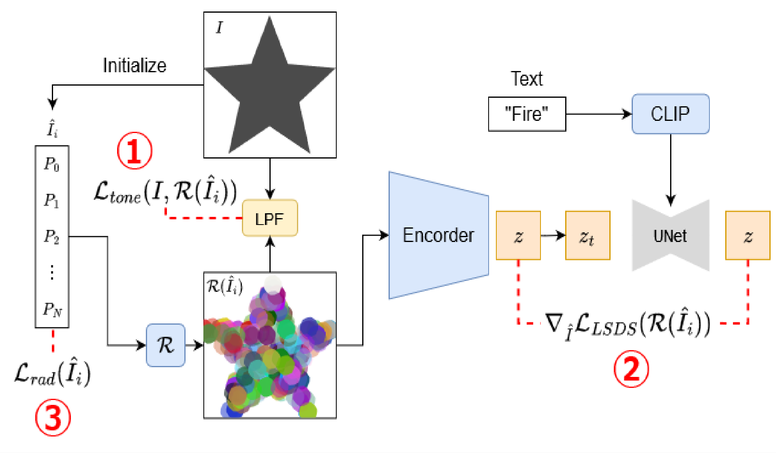
Geometric control of parameters
Place coordinate points in one direction relative to the center point

$$\mathcal{L}_{rad} = \sum_{n} \mathrm{ReLU}(\angle p_n c p_{n+1} - \angle p_n c p_{n+1})$$



Total Loss

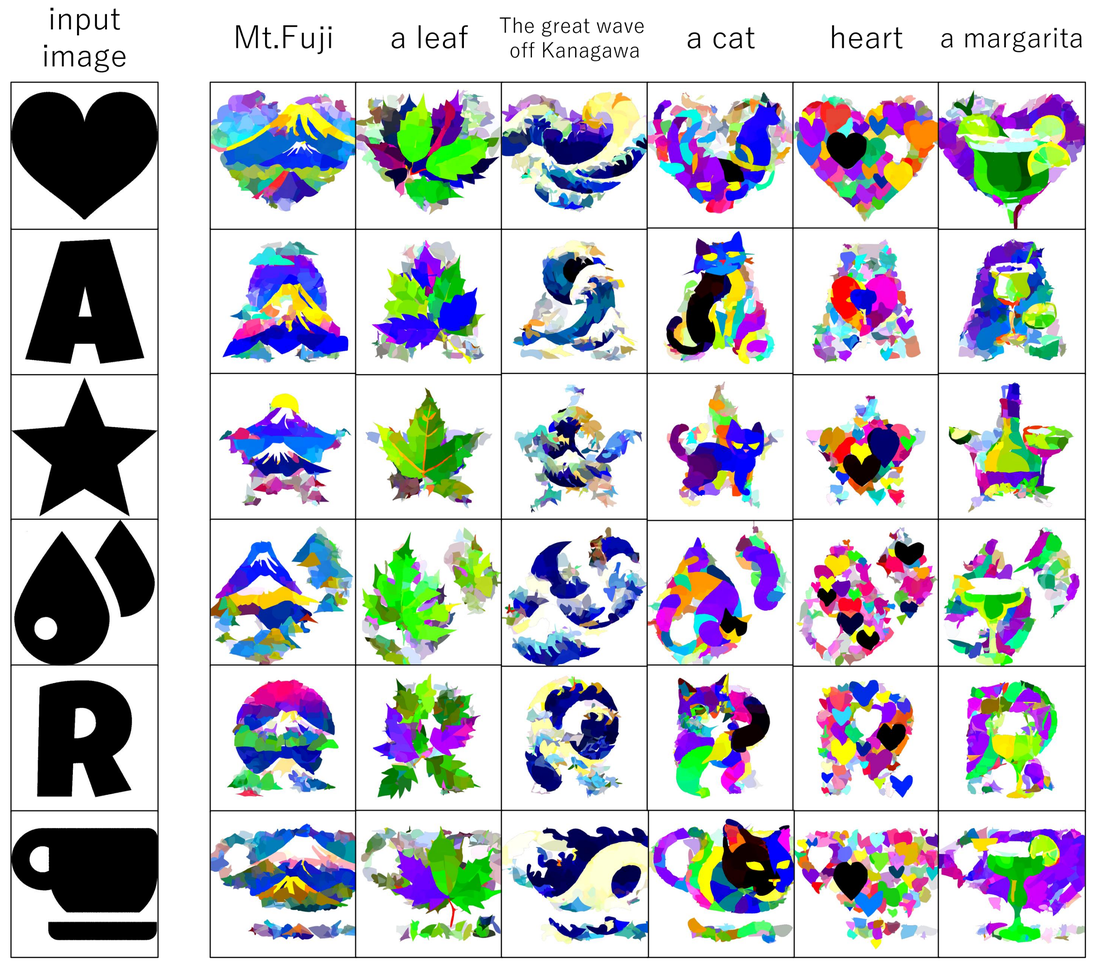
$$\mathcal{L}_{total} = \lambda_{tone}\mathcal{L}_{tone} + \lambda_{LSDS}\mathcal{L}_{LSDS} + \lambda_{rad}\mathcal{L}_{rad}$$



Experiments

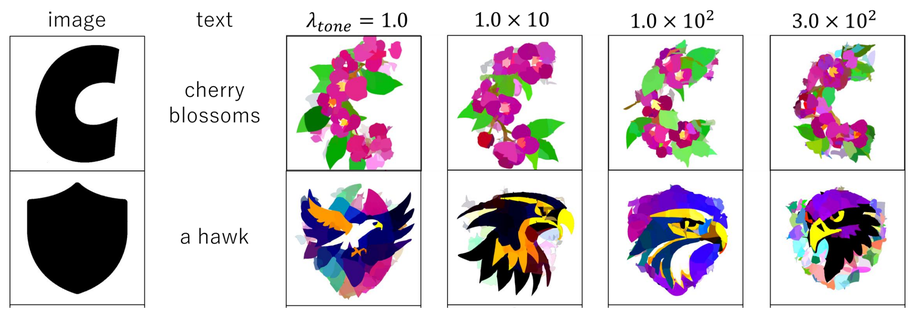
1. Generate Experiment

- Input prompts are " a logo of {concept}. minimal flat 2d vector. lineal color. trending on artstation."
- Optimize 1000 times, with $\lambda_{tone} = 200,\ \lambda_{rad} = \lambda_{LSDS} = 1$

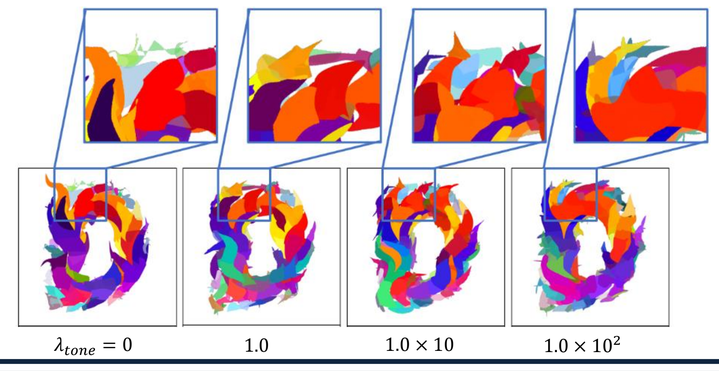


2. The case of changing loss weight

- Example of changing Tone Loss weight λ_{tone}



- Example of changing Radiation Loss weight λ_{rad}



Conclusions

- Proposes an optimization framework for generating logo images using images and text as input.
- While we were able to obtain output that met our objectives, we were unable to improve the quality of the output compared to previous research[3].

FUTURE OUTLOOK

- Manipulate the shape of the sampling result of the direct diffusion model instead of restricting the shape by region
- Further improvements to Radiation Loss

[1] Li, T., Lukáč, M., M, G. and Ragan-Kelley, J.: Differentiable Vector Graphics Rasterization for Editing and Learning, SIGGRAPH, (2020).
[2] Iluz, S., Vinker, Y., Hertz, A., Berio, D., Cohen-Or, D. and Shamir, A.: Word-As-Image for Semantic Typography, SIIGRAPH, (2023).
[3] Jain, A., Xie, A. and Abbeel, P.: VectorFusion: Text-to-SVG by Abstracting Pixel-Based Diffusion Models, CVPR,(2023).
[4] Frans, K., Soros, L. and Witkowski, O.: CLIPDraw: Exploring Text-to-Drawing Synthesis through Language-Image Encoders, NeurIPS, (2022).