Rethinking Style Transfer: From Pixels to Parameterized Brushstrokes

Dmytro Kotovenko* Matthias Wright* Arthur Heimbrecht Björn Ommer

* equal contribution

Heidelberg University, HCI/IWR, Germany

CVPR VIRTUAL JUNE 19-25

1 Introduction

Motivation

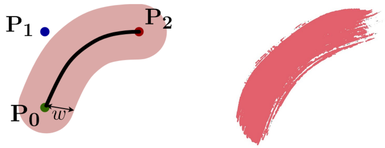
- Current Style Transfer Methods represent a stylized image as a collection of pixels
- However, artists paint with brushstrokes, not with pixels
- We propose to represent a stylized image as a collection of parameterized brushstrokes

Representation

- We define each brushstroke with a quadratic Bézier curve

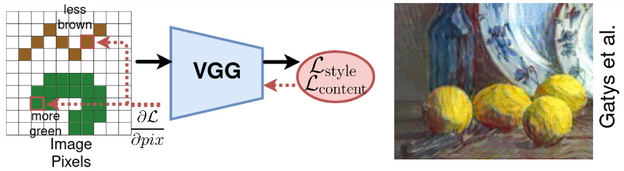
$$\mathbf{B}(t) = (1-t)^2\mathbf{P}_0 + 2(1-t)t\mathbf{P}_1 + t^2\mathbf{P}_2\,,\ 0 \le t \le 1$$

- In addition, each brushstroke has parameters for width, color, and position on the canvas

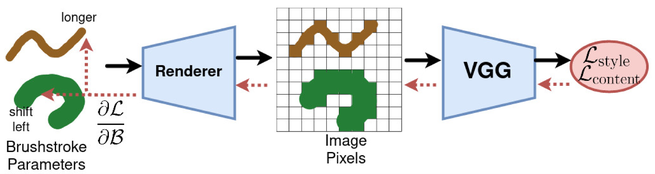


+ Position on the Canvas

2 Approach



Gatys et al. optimize individual pixels with respect to the content and style losses.



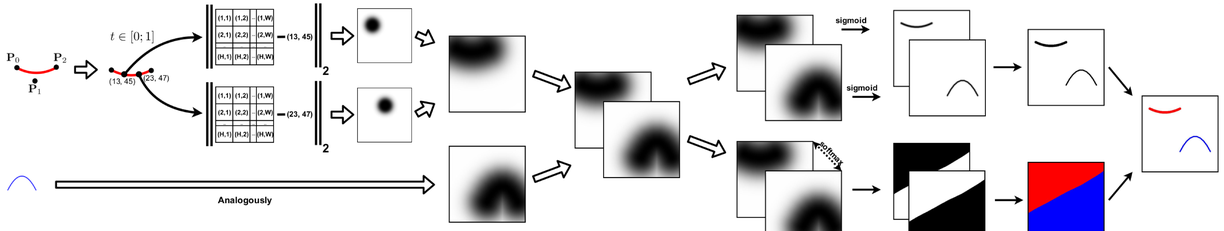
We optimize brushstroke parameters with respect to the same losses.

$$\mathcal{L}_{style} = \sum_{l=1}^{L} \frac{w_l}{N_l^2 \cdot M_l^2} ||G_r^l - G_s^l||_F$$

$$\mathcal{L}_{content} = ||\phi_l(I_r)) - \phi_l(I_c))||_2$$

G - Gram matrix of VGG activations
$\phi(\cdot)$ - VGG layer
I_r, I_s, I_c - Canvas, Content, Style

3 Differentiable Renderer



Start with Brushstroke Parameters

Sample points on a Bézier curve

Subtract sampled point's coordinates from a tensor of coordinates that represents the canvas

Compute distances from each sampled point to each pixel on the canvas

Aggregate distances from all sampled points

Repeat for every brushstroke to obtain a tensor of distances from each pixel to each brushstroke

Top: Apply sigmoid to distance tensor
Bottom: Apply softmax across brushstroke dimension to obtain the nearest brushstroke for each pixel

Top: Obtain masks for each brushstroke
Bottom: Obtain assignment matrix for nearest brushstroke

Top: Combine all masks
Bottom: Apply color to each brushstroke dimension and combine them

Compute the dot product between brushstroke mask and color assignment matrix to obtain the final rendering

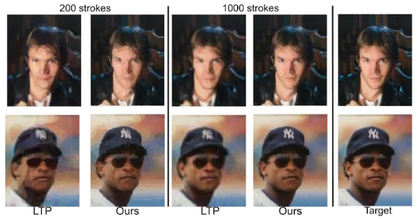
Simplified Version, see Code for details

4 Evaluation

Deception Score (DS): How similar is the stylized image to the style of an artist?
Human Deception Rate (HDR): Whether human subjects can distinguish patches of real artworks from patches of the stylization.

Method	AdaIN	WCT	Gatys et al.	AST	**Ours**	Wikiart test	Photos
Mean DS ↑	0.08	0.11	0.389	0.451	**0.588**	0.687	0.002
Mean HDR ↑	0.035	0.091	0.139	0.146	**0.268**	-	-

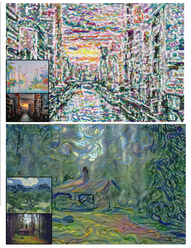
Comparison of our renderer to Learning to Paint (LTP) by Huang et al. on the image reconstruction task

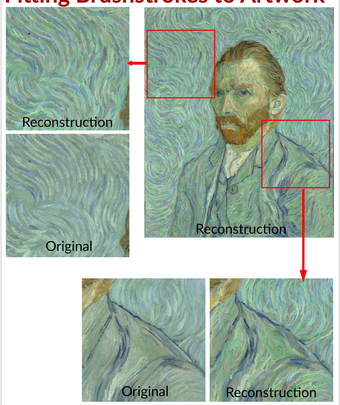


Our renderer achieves 20% lower Mean Squared Error (MSE) for 200 brushstrokes and 49% lower MSE for 1000 brushstrokes.

5 Results

Controlling the Flow of Brushstrokes



Comparisons

Fitting Brushstrokes to Artwork