Style Transfer by Rigid Alignment in Neural Net Feature Space

Suryabhan Singh Hada   Miguel Á. Carreira-Perpiñán
shada@ucmerced.edu
mcarreira-perpinan@ucmerced.edu
Dept. CSE, University of California, Merced

November 28, 2020

Style transfer problem

Content(C) | Style(S) | $f(C, S)$

- Goal: Perform style transfer in real-time for an arbitrary pair of content and style image.

Problems with current methods

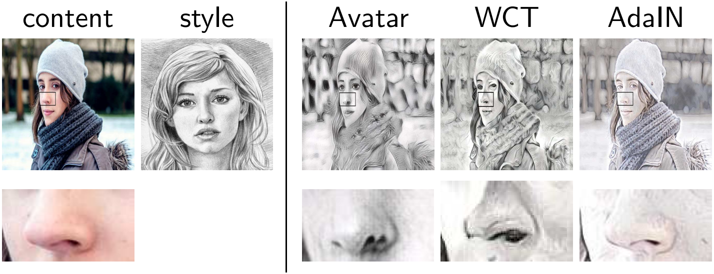


Figure: Content distortion during style transfer. Regions marked by bounding boxes are zoomed in for a better visualization.

Problems with current methods

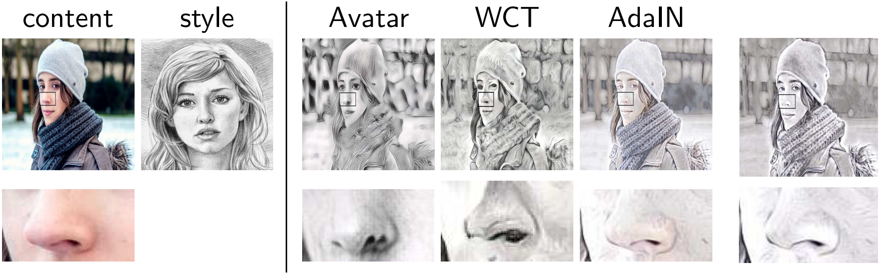


Figure: Content distortion during style transfer. Regions marked by bounding boxes are zoomed in for a better visualization.

Our hypothesis

- Working in deep neural network feature space can transfer texture and other high-level information from one image to another without altering the image structure much.
- We consider images as a collection of points in feature space, where each point represents some structural information, and if we align these points clouds using rigid alignment, we can transform these points without introducing any distortion. Here we align style features to content features.

- Matching channel-wise mean and variance of content features to those of style features. This transfers low-level style information such as colors.
- Rigid alignment:
  - Content features $\mathbf{z}_c \in \mathbb{R}^{C \times H \times W}$.
  - Style features $\mathbf{z}_s \in \mathbb{R}^{C \times H \times W}$.
  - Consider both features as point clouds of size C with each point is in $\mathbb{R}^{HW}$ space, i.e. $\mathbf{z}_c, \mathbf{z}_s \in \mathbb{R}^{C \times HW}$.

Need to arrange features in $\mathbb{R}^{C\times HW}$ space

Figure: *Third column:* no content distortion in style transfer with features $(\mathbf{z})$ transformed as C cloud points and each in $\mathbb{R}^{HW}$ space. *Fourth column:* complete content distortion in style transfer with HW cloud points and each in $\mathbb{R}^{C}$ space.

Rigid alignment

- **Step-I: Shifting:**

$$\bar{\mathbf{z}}_c = \mathbf{z}_c - \boldsymbol{\mu}_c \qquad \bar{\mathbf{z}}_s = \mathbf{z}_s - \boldsymbol{\mu}_s.$$

- **Step-II: Scaling:**

$$\hat{\mathbf{z}}_c = \frac{\bar{\mathbf{z}}_c}{\|\mathbf{z}_c\|_F} \qquad \hat{\mathbf{z}}_s = \frac{\bar{\mathbf{z}}_s}{\|\mathbf{z}_s\|_F}.$$

- **Step-III: Rotation:**

$$\arg\min_{\mathbf{Q}} \|\hat{\mathbf{z}}_s \mathbf{Q} - \hat{\mathbf{z}}_c\|_2^2 \quad \text{s.t.} \quad \mathbf{Q} \text{ is orthogonal.}$$

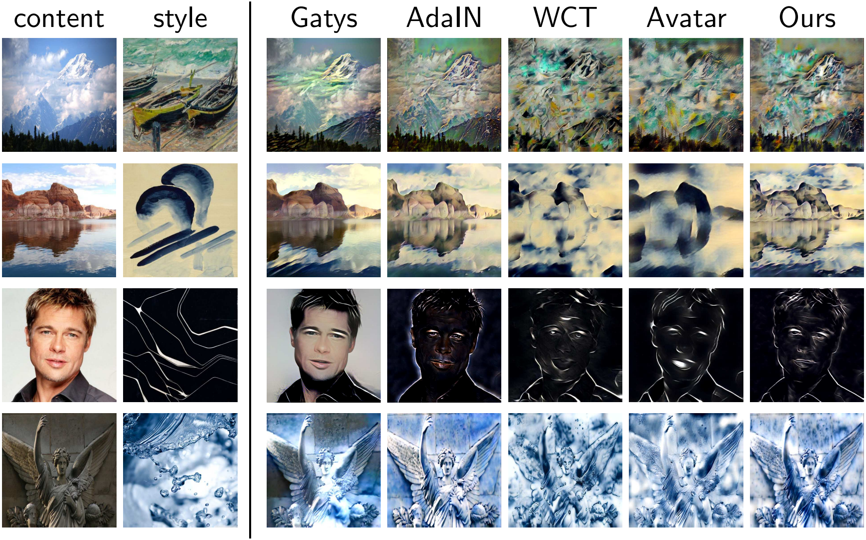
This has closed form solution: $\mathbf{Q} = \mathbf{V}\mathbf{U}^T$, where SVD$(\hat{\mathbf{z}}_c^T \hat{\mathbf{z}}_s) = \mathbf{U}\mathbf{S}\mathbf{V}^T$.

- **Step-IV: Alignment:**

$$\mathbf{z}_{sc} = \|\mathbf{z}_c\|_F \hat{\mathbf{z}}_s \mathbf{Q} + \boldsymbol{\mu}_c$$

Pass $\mathbf{z}_{sc}$ to a decoder to get the final styled image.

Some results



Conclusion

- Proposed method can achieve effective arbitrary style transfer in real-time without any content distortion.
- Does not require any training for style transfer.
- Provides a closed form solution to style transfer problem.

Thank You !