MaeSTrO: Mobile Style Transfer Orchestration using Adaptive Neural Networks

Max Reimann
Hasso Plattner Institute for Digital Engineering, University of Potsdam

Amir Semmo
Hasso Plattner Institute for Digital Engineering, University of Potsdam

Jürgen Döllner
Hasso Plattner Institute for Digital Engineering, University of Potsdam

Sebastian Pasewaldt
Digital Masterpieces GmbH, Germany

Mandy Klingbeil
Digital Masterpieces GmbH, Germany

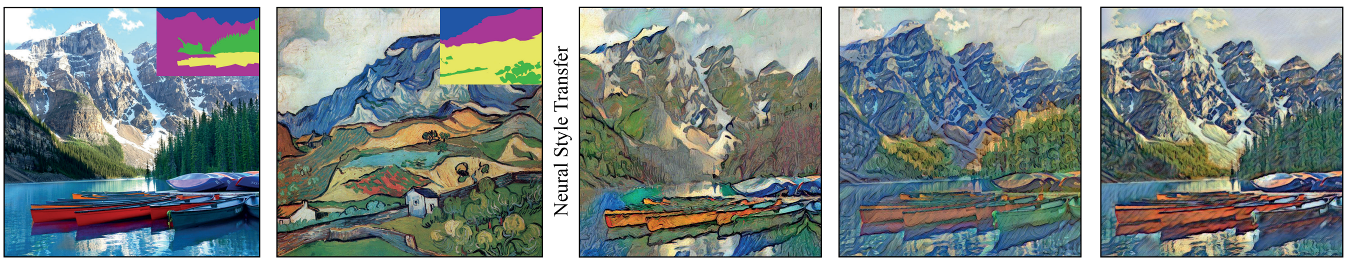


Figure 1: Comparison of two neural style transfer techniques implemented with *MaeSTrO*. Compared to the original global style transfer, the provided tools for local control (color-coded insets) are able to yield more expressive results. Content image © Matthew Fournier on Unsplash.com, used with permission.

ABSTRACT

We present *MaeSTrO*, a mobile app for image stylization that empowers users to direct, edit and perform a neural style transfer with creative control. The app uses iterative style transfer, multi-style generative and adaptive networks to compute and apply flexible yet comprehensive style models of arbitrary images at run-time. Compared to other mobile applications, *MaeSTrO* introduces an interactive user interface that empowers users to orchestrate style transfers in a two-stage process for an individual visual expression: first, initial semantic segmentation of a style image can be complemented by on-screen painting to direct sub-styles in a spatially-aware manner. Second, semantic masks can be virtually drawn on top of a content image to adjust neural activations within local image regions, and thus direct the transfer of learned sub-styles. This way, the general feed-forward neural style transfer is evolved towards an interactive tool that is able to consider composition variables and mechanisms of general artwork production, such as color, size and location-based filtering. *MaeSTrO* additionally enables users to define new styles directly on a device and synthesize high-quality images based on prior segmentations via a service-based implementation of compute-intensive iterative style transfer techniques.

SIGGRAPH '18 Appy Hour, August 12-16, 2018, Vancouver, BC, Canada

ACM ISBN 978-1-4503-5807-1/18/08.
https://doi.org/10.1145/3213779.3213783

CCS CONCEPTS

• **Computing methodologies** → **Non-photorealistic rendering**; **Image processing**;

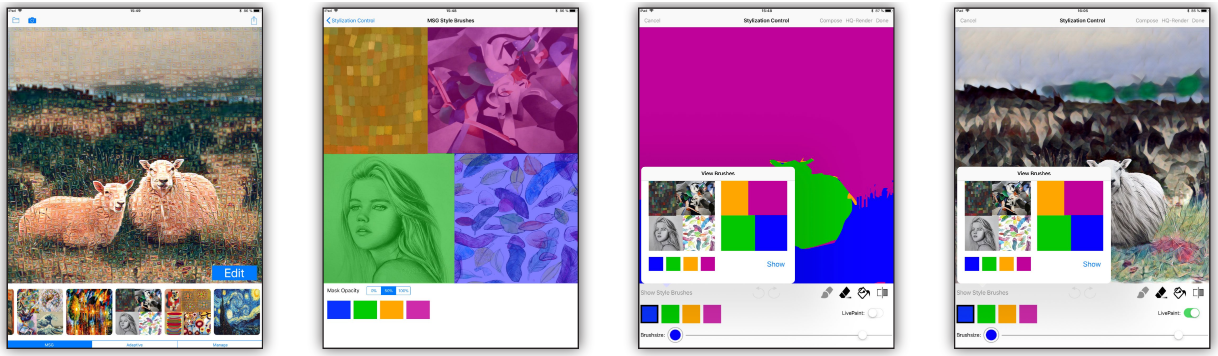
KEYWORDS

neural style transfer, mobile devices, artistic rendering, interaction

ACM Reference Format:
Max Reimann, Amir Semmo, Jürgen Döllner, Sebastian Pasewaldt, and Mandy Klingbeil. 2018. MaeSTrO: Mobile Style Transfer Orchestration using Adaptive Neural Networks. In *Proceedings of SIGGRAPH '18 Appy Hour*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3213779.3213783

1 MOTIVATION

Image filters, particularly those used for mobile expressive rendering, have become a pervasive technology for casual creativity and users that seek unique possibilities to stylize images [Dev 2013]. For instance, mobile artists—a new user group of serious hobbyists with high standards—are eager to adapt to powerful and flexible tools that facilitate their creative work. Image filters are traditionally implemented by following an engineering approach; providing low- and high-level control over the stylization process. With the advent of neural style transfer technology [Gatys et al. 2016], mobile image filtering apps have increasingly evolved into "one-click solutions" that allow to transfer a pre-defined style image to a content image (Figure 1). Although this approach enables to easily create artistic renditions—without having prior knowledge of photo-manipulation software—the underlying technology faces inherent limitations regarding low-level control for localized image stylization [Semmo et al. 2017a], hindering creative control over the results.



Global Stylization | Style Image Masks | Content Image Masks | Live-Painting Screen

Figure 2: Screenshots of *MaeSTrO*: Global stylization can be refined by defining *content image masks* in the *live-painting screen*. A color-mapping is used to ease the mapping between style and content image masks. Content image © Rick Barrett on Unsplash.com, used with permission.

In this work, we present *MaeSTrO*, an iOS app that implements and enhances style transfer technologies to allow for local creative control that facilitates an interactive, artistic image editing. Our app targets mobile artists with basic image editing know-how by using established on-screen painting metaphors for the local definition of sub-styles and the successive application to content images.

2 TECHNICAL APPROACH

MaeSTrO implements three different neural network techniques, each providing a trade-off between usability and picture quality. Single-style feed-forward [Johnson et al. 2016] are currently used in the majority of techniques for mobile style transfer (e.g., [Semmo et al. 2017b]), since they enable nearly interactive performance, even on mobile devices. Once trained off-line, the feed-forward network—representing a single style—is globally applied to the whole input image. To cope with this limitation while maintaining a short computation time, multi-style generative networks (*MSG-Net*) are utilized and extended [Zhang and Dana 2017]. Using semantic masks for style images, these networks can be trained on multiple style images and enable local style-blending in feature space, yielding smooth transitions between multiple styles. Although *MSG-Nets* improve creative control, users are still limited to apply pre-trained styles. To enable an on-device style definition, *MaeSTrO* additionally implements the approach of Huang and Belongie [2017] that performs a style transfer for arbitrary styles defined on-device by using an encoder-decoder network containing an adaptive instance normalization (*adaIn*). Similar to the MSG-Net approach, we extended the *adaIn-network* by semantic masks to allow for local control of style definitions and applications. Also the third technique, the iterative style transfer approach [Gatys et al. 2016] implements local control through segmentation masks [Luan et al. 2017] and enables the application of arbitrary styles. However, the computational complexity of the approach does not enable an on-device application. Thus, it is implemented as a web service, where users can define and modify styles on a mobile device, for example using the *adaIn* approach, and request the web service to perform the high-quality style transfer.

All implemented approaches enable local control of the style application to a content image. In addition, the *adaIN* and iterative approaches enable users to define sub-styles, i.e., locally constrained regions that are assigned to different styles (Figure 2). The definition and application of sub-styles is implemented using pixel-precise painting metaphors. When editing a content image, an overlay provides additional information about which sub-style is mapped to which virtual brush.

The iterative approach is implemented using PyTorch and the on-device approaches are implemented using CoreML for the iOS operating system. The style transfer run-time performance depends on the number of sub-styles applied as well as of the image resolution. For example, the application of two sub-styles for an 720×720 image takes approx. 1.0 second for *adaIn* and 1.5 seconds for *MSG* on an iPad Pro 10.5". To allow for interactive mask application, a live painting mode has been implemented that directly shows the application of pre-computed sub-styles, while the final image synthesis is performed afterwards.

ACKNOWLEDGEMENTS

This work was funded by the Federal Ministry of Education and Research (BMBF), Germany, for the AVA project 01IS15041.

REFERENCES

Kapil Dev. 2013. Mobile Expressive Renderings: The State of the Art. *IEEE Computer Graphics and Applications* 33, 3 (May/June 2013), 22–31. https://doi.org/10.1109/MCG.2013.20

Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. 2016. Image Style Transfer Using Convolutional Neural Networks. In *Proc. CVPR*. IEEE Computer Society, Los Alamitos, 2414–2423. https://doi.org/10.1109/CVPR.2016.265

Xun Huang and Serge Belongie. 2017. *Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization.* arXiv.org report 1703.06868. arXiv. https://arxiv.org/abs/1703.06868

Justin Johnson, Alexandre Alahi, and Li Fei-Fei. 2016. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In *Proc. ECCV*. Springer International, Cham, Switzerland, 694–711. https://doi.org/10.1007/978-3-319-46475-6_43

Fujun Luan, Sylvain Paris, Eli Shechtman, and Kavita Bala. 2017. *Deep Photo Style Transfer.* CoRR abs/1703.07511. arXiv. http://arxiv.org/abs/1703.07511

Amir Semmo, Tobias Isenberg, and Jürgen Döllner. 2017a. Neural Style Transfer: A Paradigm Shift for Image-based Artistic Rendering?. In *Proc. NPAR*, Holger Winnemöller and Lyn Bartram (Eds.). ACM, New York, 5:1–5:13. https://doi.org/10.1145/3092919.3092920

Amir Semmo, Matthias Trapp, Jürgen Döllner, and Mandy Klingbeil. 2017b. Pictory: Combining Neural Style Transfer and Image Filtering. In *Proc. SIGGRAPH Appy Hour*. ACM, New York, NY, USA, 5:1–5:2. https://doi.org/10.1145/3098900.3098906

Hang Zhang and Kristin Dana. 2017. *Multi-style Generative Network for Real-time Transfer.* arXiv.org report 1703.06953. arXiv. https://arxiv.org/abs/1703.06953