StyleTune: Interactive Style Transfer Enhancement on Mobile Devices

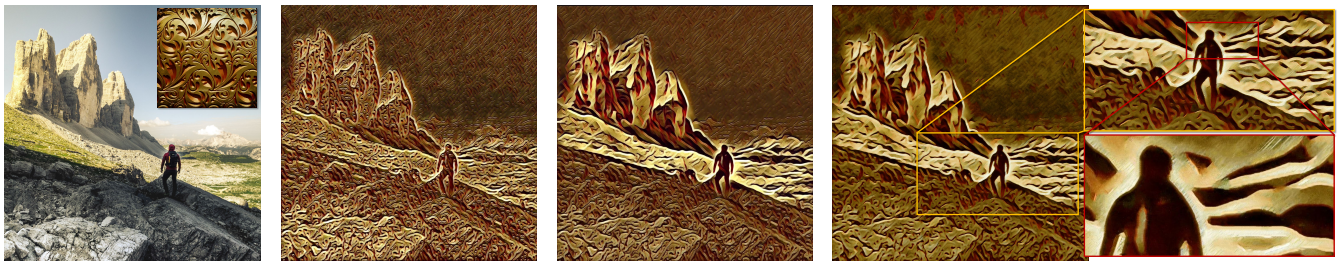
Benito Buchheim
Hasso Plattner Institute for Digital Engineering, University of Potsdam
Germany

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

Sebastian Pasewaldt
Digital Masterpieces GmbH
Germany

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

Matthias Trapp
Hasso Plattner Institute for Digital Engineering, University of Potsdam
Germany



(a) Content and Style (b) Global Transfer (c) Local Adjustments (d) Upsampled NST and zoomed versions

Figure 1: Comparison of different NST results that have been created with *StyleTune*. Besides a global transfer, a user can also apply local adjustments. To enable high quality and high resolution output, we implement a patch-based upsampled NST. Content image by Maximilian Pawlikowsky © Shutterstock, used with permission.

ABSTRACT

We present *StyleTune*, a mobile app for interactive style transfer enhancement that enables global and spatial control over stroke elements and can generate high fidelity outputs. The app uses adjustable neural style transfer (NST) networks to enable art-direction of stroke size and orientation in the output image. The implemented approach enables continuous and seamless edits through a unified stroke-size representation in the feature space of the style transfer network. *StyleTune* introduces a three-stage user interface, that enables users to first explore global stroke parametrizations for a chosen NST. They can then interactively locally retouch the stroke size and orientation using brush metaphors. Finally, high resolution outputs of 20 Megapixels and more can be obtained using a patch-based upsampling and local detail transfer approach, that transfers small-scale details such as paint-bristles and canvas structure. The app uses Apple's CoreML and Metal APIs for efficient on-device processing.

SIGGRAPH '21 Appy Hour, August 09-13, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8358-5/21/08.
https://doi.org/10.1145/3450415.3464400

CCS CONCEPTS

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

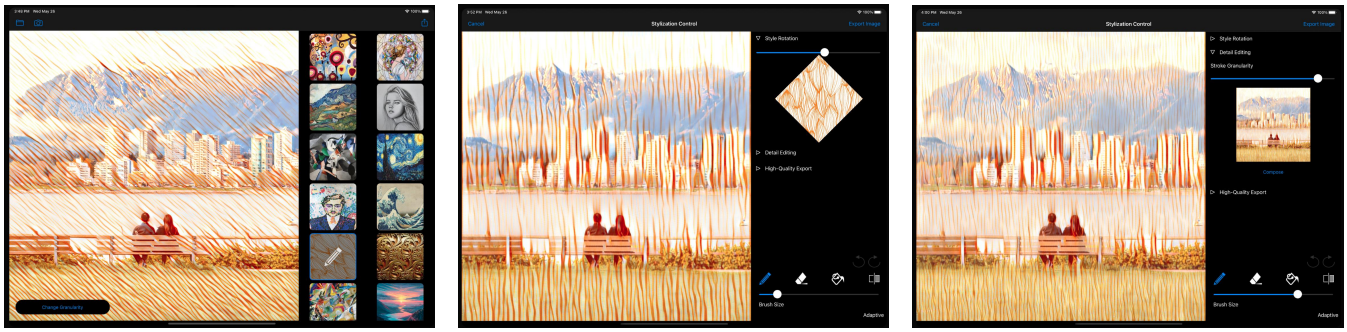
KEYWORDS

neural style transfer, mobile devices, artistic rendering, interaction

ACM Reference Format:
Benito Buchheim, Max Reimann, Sebastian Pasewaldt, Jürgen Döllner, and Matthias Trapp. 2021. StyleTune: Interactive Style Transfer Enhancement on Mobile Devices. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Appy Hour (SIGGRAPH '21 Appy Hour), August 09-13, 2021.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3450415.3464400

1 MOTIVATION

With the advancement of mobile processing capabilities, an increasing number of mobile apps use image filtering techniques to enable users to artistically stylize images and foster casual creativity. In recent years, neural style transfer, introduced by Gatys et al. [2016], has received a lot of attention in both research and application as a generalized, learning-based method to transfer the style from an exemplar image to a content image. Feed-forward style transfer networks [Johnson et al. 2016] have furthermore gained popularity in several mobile apps for non-professional image filtering and editing, such as Prisma. The majority of these apps, however, implement NSTs in form of a one-click solution which does not allow the user to alter results or make artistic decisions. Becasso [Pasewaldt et al.



(a) Style selection (b) Global orientation adjustment (c) Local retouching

Figure 2: Screenshots of *StyleTune*: After selecting a style (a)the stroke orientation can be adjusted (b). The user can then locally retouch strokes in the image using *brush metaphors* (c). Content image by karamysh © Shutterstock, used with permission.

2016], introduces a form of artistic control by combining NSTs with a post-processing step using image–based artistic rendering, while MaeSTrO [Reimann et al. 2019] combines different pretrained styles in a single output image. Compared to previous mobile approaches, our app *StyleTune* enhances the level of control over the output by enabling the user to interactively alter individual aspects of the style such as stroke orientation and size.

2 TECHNICAL APPROACH

StyleTune implements stroke–adjustable network architectures that can globally and locally vary the stroke size in the output image. Two adjustable network variants are implemented in the app. The first variant is based on the architecture proposed by Jing et al. [2018] and encodes discrete style sizes in different network branches that are blended together according to a weighting factor before decoding the output image. During on-device execution, the stroke size can be locally applied either using pixel-space blending of multiple stroke-scales for quick previews, or blending features from different network branches and jointly decoding these to create seamless transitions between brushstrokes. This approach works especially well for making small-scale local edits. Furthermore, as a second variant, we extend a feed-forward style transfer network architecture by Johnson et al. [2016] with a second branch for control over style element size using adaptive upscaling. This approach works especially well for large-scale adjustments on the global level, as the structure of style elements is preserved on all scales. Stroke orientation edits in *StyleTune* are implemented by making use of the rotation variance of convolutional networks. Rotating the input by a certain amount and then rotating the style transfer output back to its original state thus results in a orientation change of the strokes, without requiring changes to the model architecture or training.

Limited mobile device memory is a constraint on the achievable output size of mobile NST, resulting in output sizes limited to a few megapixels, which is usually considerably less than the input resolution captured by the smartphone camera. To overcome this issue and enable the generation of high resolution NST images, we make use of a patch-based upsampling technique introduced by Texler et al. [2019]. Given a very high resolution style image, it upsamples the previously obtained low resolution NST result to dimensions of 20 Megapixels and more, while transferring fine details such as paint bristles or the canvas texture from the style image.

The interactive editing and enhancement workflow in *StyleTune* is as follows. First the user selects and applies a style to the content (Figure 2a), after which he is presented with global style settings, where style granularity and orientation can be adjusted using sliders (Figure 2b). Using painting brush metaphors, the users can then locally apply different stroke-size and orientation edits. Strokes are blended in image space to retain interactivity in the drawing process and can be merged on demand to create seamless stroke transitions (Figure 2c). In a final step, the composition can be exported at a very high resolution using patch-based upsampling. The app implementation is based on Apple's CoreML and Metal APIs for GPU-based processing of neural networks and rendering techniques.

ACKNOWLEDGMENTS

This work was partially funded by the German Federal Ministry of Education and Research (BMBF) through grants 01IS18092 ("mdViPro") and 01IS19006 ("KI-Labor ITSE")

REFERENCES

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

Yongcheng Jing, Yang Liu, Yezhou Yang, Zunlei Feng, Yizhou Yu, and Mingli Song. 2018. Stroke Controllable Fast Style Transfer with Adaptive Receptive Fields. *CoRR* abs/1802.07101 (2018), 244–260. arXiv:1802.07101 http://arxiv.org/abs/1802.07101

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

Sebastian Pasewaldt, Amir Semmo, Jürgen Döllner, and Frank Schlegel. 2016. Becasso: Artistic Image Processing and Editing on Mobile Devices. In *Proc. MGIA* (Macau). ACM, New York, NY, USA, 14:1–14:1. https://doi.org/10.1145/2999508.2999518

Max Reimann, Mandy Klingbeil, Sebastian Pasewaldt, Amir Semmo, Matthias Trapp, and Jürgen Döllner. 2019. Locally controllable neural style transfer on mobile devices. *The Visual Computer* 35, 11 (2019), 1531–1547. https://doi.org/10.1007/s00371-019-01654-1

Ondřej Texler, Jakub Fišer, Michal Lukáč, Jingwan Lu, Eli Shechtman, and Daniel Sýkora. 2019. Enhancing Neural Style Transfer using Patch-Based Synthesis. In *Proceedings of the 8th ACM/EG Expressive Symposium*. Eurographics Association, Goslar, DEU, 43–50.