Pictory: Combining Neural Style Transfer and Image Filtering

Amir Semmo Matthias Trapp Jürgen Döllner
Hasso-Plattner-Institut,
Faculty of Digital Engineering,
University of Potsdam, Germany
{first.last}@hpi.uni-potsdam.de

Mandy Klingbeil
Hasso-Plattner-Institut,
University of Potsdam, Germany
Digital Masterpieces GmbH, Germany
mandy.klingbeil@hpi.uni-potsdam.de

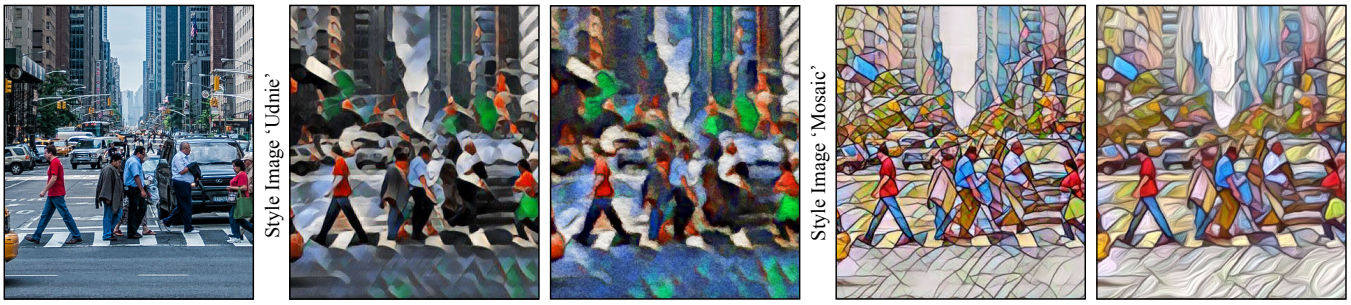

Figure 1: Outputs of the mobile app *Pictory* that combines the results of a feed-forward NST [Johnson et al. 2016] with image filtering to inject paint characteristics (here: watercolor, oil paint). Content image by Frank Köhntopp is in the public domain.

ABSTRACT

This work presents *Pictory*, a mobile app that empowers users to transform photos into artistic renditions by using a combination of neural style transfer with user-controlled state-of-the-art nonlinear image filtering. The combined approach features merits of both artistic rendering paradigms: deep convolutional neural networks can be used to transfer style characteristics at a global scale, while image filtering is able to simulate phenomena of artistic media at a local scale. Thereby, the proposed app implements an interactive two-stage process: first, style presets based on pre-trained feed-forward neural networks are applied using GPU-accelerated compute shaders to obtain initial results. Second, the intermediate output is stylized via oil paint, watercolor, or toon filtering to inject characteristics of traditional painting media such as pigment dispersion (watercolor) as well as soft color blendings (oil paint), and to filter artifacts such as fine-scale noise. Finally, on-screen painting facilitates pixel-precise creative control over the filtering stage, e. g., to vary the brush and color transfer, while joint bilateral upsampling enables outputs at full image resolution suited for printing on real canvas.

SIGGRAPH '17 Appy Hour, Los Angeles, CA, USA

CCS CONCEPTS

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

KEYWORDS

mobile, neural style transfer, image filtering, artistic rendering

ACM Reference format:
Amir Semmo, Matthias Trapp, Jürgen Döllner, and Mandy Klingbeil. 2017. Pictory: Combining Neural Style Transfer and Image Filtering In *Proceedings of SIGGRAPH '17 Appy Hour, Los Angeles, CA, USA, July 30 - August 03, 2017,* 2 pages.
DOI: 10.1145/3098900.3098906

1 MOTIVATION

Image-based artistic rendering (IB-AR) enjoys a growing popularity in mobile expressive rendering [Dev 2013; Winnemöller 2013] to simulate the appeal of traditional artistic styles and media for visual communication [Kyprianidis et al. 2013; Rosin and Collomosse 2013] such as oil paint, watercolor, and cartoon. Classical IB-AR paradigms typically simulate their characteristics and phenomena by a feature-level engineering approach, e. g., to locally direct the smoothing and adjustment of image colors via filtering. A more generalized approach has been introduced by the architecture engineering approach of deep learning, which activates layers of pre-trained deep convolutional neural networks (CNNs) to match content and style statistics, and thus perform a neural style transfer (NST) between arbitrary images [Gatys et al. 2016]. While first applications demonstrate the practicability of NSTs by the example of color and texture transfers as well as casual creativity apps (e. g., Prisma), local effects and phenomena of traditional artistic media at high-fidelity and resolution are still hard to reproduce.

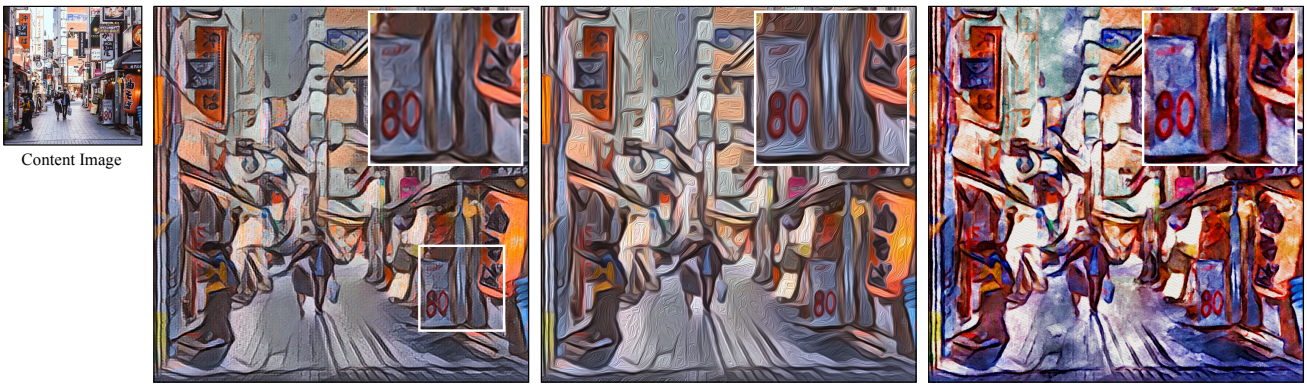

Figure 2: Results produced for an input image with a resolution of $2,048 \times 2,048$ pixels. The low-resolution NST result (512×512 pixels) is used with the high-resolution input for flow-based joint bilateral upsampling (FJBU). Afterward, post-process image filtering is performed to locally inject paint characteristics. Content image by Redd Angelo is in the public domain.

We conjecture that NSTs may be used as one of multiple processing stages and combined with the knowledge and algorithms of other paradigms [Semmo et al. 2017]. NSTs would thus operate as a first stage of image processing to introduce higher-level abstractions—to be followed by low-level, established filtering techniques to simulate drawing media and, e. g., their interplay with substrates (Figure 1).

2 TECHNICAL APPROACH

This work presents *Pictory*, a mobile app that combines NSTs with image filtering. At this, the generative approach of Johnson et al. [Johnson et al. 2016] is combined with the image processing framework of Semmo et al. [Semmo et al. 2016] to implement interactive filtering. Thereby, image abstraction at a global scale is combined with local paint effects such as edge darkening, pigment density variation, and wet-in-wet of watercolor [Bousseau et al. 2006; Wang et al. 2014], and smooth continuous oilpaint-like texture effects via flow-based Gaussian filtering with Phong shading [Hertzmann 2002; Semmo et al. 2016]. Figure 2 shows an output where the abstract style of Pablo Picasso's "La Muse" is used to generate an effect of higher-level abstraction, before adding mentioned filters to inject the respective low-level paint characteristics. Each of the filtering effects can be locally parameterized by image masking, e. g., over the color and texture transfer modality of the NST or the filters' parameters such as wetness, smoothness, and relief.

The mobile app was implemented using the OpenGL ES Shading Language using compute shaders, and was deployed on Android. To process images with full HD resolution, neural networks with reduced layers for the convolutional stages are applied in a tile-based approach to optimize processing time and memory consumption. In addition, flow-based joint bilateral upsampling [Kopf et al. 2007; Semmo et al. 2016] of the low-resolution NST result is performed with the high-resolution input to reduce visual noise and obtain fine paint structures at the filtering stage (Figure 2). Using these optimizations, our app provides initial NST results between 2 seconds (512×512 pixels) and 10 seconds (1024×1024 pixels), and enables post-process filtering at interactive frame rates on a Google™ Pixel C with a NVIDIA® Maxwell 256 core GPU.

ACKNOWLEDGMENTS

We would like to thank Moritz Hilscher and Hendrik Tjabben for their substantial contributions to the app prototype. This work was funded by the Federal Ministry of Education and Research (BMBF), Germany, for the AVA project 01IS15041B and within the InnoProfile Transfer research group "4D*n*D-Vis" (www.4dndvis.de).

REFERENCES

Adrien Bousseau, Matt Kaplan, Joëlle Thollot, and François X. Sillion. 2006. Interactive Watercolor Rendering with Temporal Coherence and Abstraction. In *Proc. NPAR*. ACM, New York, 141–149. doi: 10.1145/1124728.1124751

Kapil Dev. 2013. Mobile Expressive Renderings: The State of the Art. *IEEE Computer Graphics and Applications* 33, 3 (May/June 2013), 22–31. doi: 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. doi: 10.1109/CVPR.2016.265

Aaron Hertzmann. 2002. Fast Paint Texture. In *Proc. NPAR*. ACM, New York, 91–96. doi: 10.1145/508530.508546

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. doi: 10.1007/978-3-319-46475-6_43

Johannes Kopf, Michael F. Cohen, Dani Lischinski, and Matt Uyttendaele. 2007. Joint Bilateral Upsampling. *ACM Transactions on Graphics* 26, 3, Article 96 (July 2007). doi: 10.1145/1276377.1276497

Jan Eric Kyprianidis, John Collomosse, Tinghuai Wang, and Tobias Isenberg. 2013. State of the "Art": A Taxonomy of Artistic Stylization Techniques for Images and Video. *IEEE Transactions on Visualization and Computer Graphics* 19, 5 (May 2013), 866–885. doi: 10.1109/TVCG.2012.160

Paul Rosin and John Collomosse (Eds.). 2013. *Image and Video based Artistic Stylisation*. Computational Imaging and Vision, Vol. 42. Springer, London/Heidelberg. doi: 10.1007/978-1-4471-4519-6

Amir Semmo, Tobias Dürschmid, Matthias Trapp, Mandy Klingbeil, Jürgen Döllner, and Sebastian Pasewaldt. 2016. Interactive Image Filtering with Multiple Levels-of-control on Mobile Devices. In *Proc. MGIA*. ACM, New York, Article 2, 8 pages. doi: 10.1145/2999508.2999521

Amir Semmo, Tobias Isenberg, and Jürgen Döllner. 2017. Neural Style Transfer: A Paradigm Shift for Image-based Artistic Rendering?. In *Proc. NPAR*. ACM, New York. To appear.

Amir Semmo, Matthias Trapp, Tobias Dürschmid, Jürgen Döllner, and Sebastian Pasewaldt. 2016. Interactive Multi-scale Oil Paint Filtering on Mobile Devices. In *Proc. ACM SIGGRAPH Posters*. ACM, New York, 42:1–42:2. doi: 10.1145/2945078.2945120

Miaoyi Wang, Bin Wang, Yun Fei, Kanglai Qian, Wenping Wang, Jiating Chen, and Jun-Hai Yong. 2014. Towards Photo Watercolorization with Artistic Verisimilitude. *IEEE Transactions on Visualization and Computer Graphics* 20, 10 (Feb. 2014), 1451–1460. doi: 10.1109/TVCG.2014.2303984

Holger Winnemöller. 2013. NPR in the Wild. In *Image and Video based Artistic Stylisation*, Paul Rosin and John Collomosse (Eds.). Computational Imaging and Vision, Vol. 42. Springer, Chapter 17, 353–374. doi: 10.1007/978-1-4471-4519-6_17