Trios: Stylistic Rendering of 3D Photos

Ulrike Bath
ulrike.bath@student.hpi.uni-potsdam.de
Hasso Plattner Institute,
University of Potsdam
Germany

Sumit Shekhar
sumit.shekhar@hpi.uni-potsdam.de
Hasso Plattner Institute,
University of Potsdam
Germany

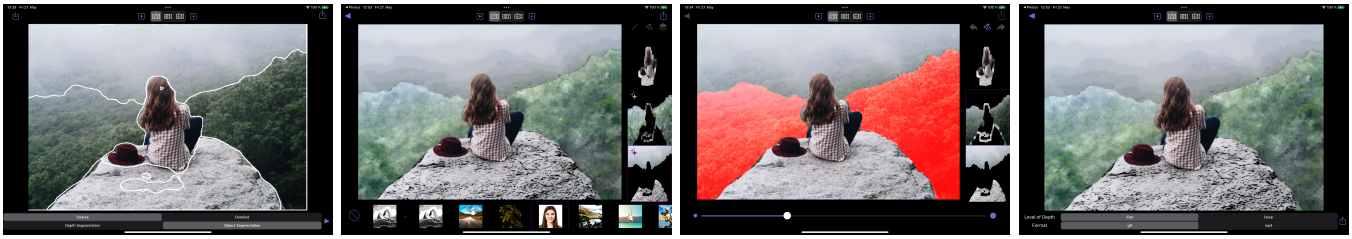
Hendrik Tjabben
tjabben@adobe.com
Adobe Systems Engineering GmbH
Germany

Amir Semmo
amir.semmo@digitalmasterpieces.com
Digital Masterpieces GmbH
Germany

Sebastian Pasewaldt
sebastian.pasewaldt@digitalmasterpieces.com
Digital Masterpieces GmbH
Germany

Jürgen Döllner
juergen.doellner@hpi.uni-potsdam.de
Hasso Plattner Institute,
University of Potsdam
Germany

Matthias Trapp
matthias.trapp@hpi.uni-potsdam.de
Hasso Plattner Institute,
University of Potsdam
Germany



(a) Segmentation Screen (b) Stylization Screen (c) Touch-Up Screen (d) Export Screen

Figure 1: Overview of the main user interface screens provided by our prototypical application.

ABSTRACT

3D photography has emerged as a medium that provides an immersive dimension to 2D photos. We present *Trios*, an interactive mobile app that combines the vividness of image-based artistic rendering with 3D photos by implementing an end-to-end pipeline for their generation and stylization. *Trios* uses Apple's accelerated image-processing APIs and dedicated Neural Engine for depth-generation and learning-based artistic rendering. The pipeline runs at interactive frame rates and outputs a compact video, which can easily be shared. Thus, it serves as a unique interactive tool for digital artists interested in creating immersive artistic content.

SIGGRAPH '22 Appy Hour, August 07-11, 2022, Vancouver, BC, Canada

ACM ISBN 978-1-4503-9365-2/22/08.
https://doi.org/10.1145/3532723.3535467

CCS CONCEPTS

• **Computing methodologies** → **Image-based rendering**; **Image processing**; **Non-photorealistic rendering**; • **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**.

KEYWORDS

3D photos, image stylization, mobile devices

ACM Reference Format:
Ulrike Bath, Sumit Shekhar, Hendrik Tjabben, Amir Semmo, Sebastian Pasewaldt, Jürgen Döllner, and Matthias Trapp. 2022. Trios: Stylistic Rendering of 3D Photos. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Appy Hour (SIGGRAPH '22 Appy Hour), August 07-11, 2022*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3532723.3535467

1 MOTIVATION

Due to the progress in mobile graphics, smartphones and tablets nowadays allow for a large range of on-device image-processing.

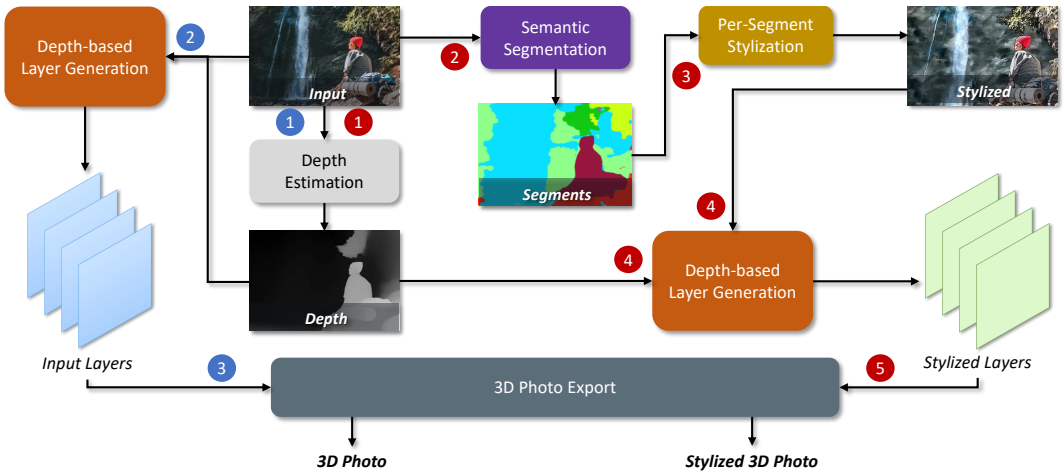


Figure 2: Schematic overview of the processing pipeline implemented in *Trios*; steps 1-3 depict the generation of a *3D Photo* and steps 1-5 the generation of a *Stylized 3D Photo*.

On-device 3D photo generation for visual immersiveness is a compelling example in this regard [Kopf et al. 2020]. Previously, apps for Image-based Artistic Rendering (IB-AR) also exploited the advanced visual computing capabilities of mobile devices [Pasewaldt et al. 2016; Semmo and Pasewaldt 2020]. As compared to traditional IB-AR approaches, 3D photos offer new possibilities with respect to stylization and visualization. In this work, we extend the visual-richness of IB-AR techniques to 3D photos. Thus, we develop a mobile-based framework for generation and stylization of 3D photos given an input RGB-image.

2 BACKGROUND & TECHNICAL OVERVIEW

Hedman *et al.* proposed the first approach for 3D photo acquisition [2017]. Since then, significant progress advanced to method being capable of running fully on a high-end smartphone [Hedman and Kopf 2018; Kopf et al. 2020; Shih et al. 2020]. However, most of the existing methods do not consider user interactivity during creation and editing of 3D photos. Our framework adopts the following approach for generation and interactive stylization of 3D photos on high-end mobile devices (Fig. 2).

Depth Estimation We employ the mobile-version of MiDaS, a depth estimation technique by Ranftl *et al.* [2022], to generate a depth map for the given input image.

Layer Generation The depth map is used to decompose the RGB image into *input layers* at different depth levels and represent these as a Layered Depth Image (LDI) [Shade et al. 1998], stored in a 2D texture array for GPU-based processing.

Semantic Segmentation The input image is separated into semantic segments for the purpose of stylization.

Per-Segment Stylization Following the above, we employ a variety of classical and learning-based IB-AR techniques for stylizing individual segments.

Stylized Layers Similar to *input layers*, the depth map is used to decompose the stylized image into *stylized layers* and is represented as an LDI.

3D Photo Export The LDI representation of *input* or *stylized* layers can be traversed in different ways using a virtual camera to obtain a 3D Photo as a video or GIF.

Our prototypical framework is based on iOS and iPadOS using Swift, UIKit, CoreImage, CoreML, and Metal Application Programming Interfaces (APIs). The principal methodology is not limited to Apple devices and can be extended for other environments as well.

ACKNOWLEDGMENTS

This work was partially funded by the BMBF (01IS18092-"mdViPro" and 01IS19006-"KI-LAB-ITSE") and the Research School on "Service-Oriented Systems Engineering" of the Hasso Plattner Institute.

REFERENCES

Peter Hedman, Suhib Alsisan, Richard Szeliski, and Johannes Kopf. 2017. Casual 3D Photography. *ACM Trans. Graph.* 36, 6, Article 234 (2017), 15 pages. https://doi.org/10.1145/3130800.3130828

Peter Hedman and Johannes Kopf. 2018. Instant 3D Photography. *ACM Trans. Graph.* 37, 4, Article 101 (jul 2018), 12 pages. https://doi.org/10.1145/3197517.3201384

Johannes Kopf, Kevin Matzen, Suhib Alsisan, Ocean Quigley, Francis Ge, Yangming Chong, Josh Patterson, Jan-Michael Frahm, Shu Wu, Matthew Yu, Peizhao Zhang, Zijian He, Peter Vajda, Ayush Saraf, and Michael Cohen. 2020. One Shot 3D Photography. *ACM Trans. Graph.* 39, 4, Article 76 (jul 2020), 13 pages. https://doi.org/10.1145/3386569.3392420

Sebastian Pasewaldt, Amir Semmo, Jürgen Döllner, and Frank Schlegel. 2016. Becasso: Artistic Image Processing and Editing on Mobile Devices. In *SIGGRAPH ASIA 2016 Mobile Graphics and Interactive Applications (SA '16)*. Article 14, 1 pages. https://doi.org/10.1145/2999508.2999518

René Ranftl, Katrin Lasinger, David Hafner, Konrad Schindler, and Vladlen Koltun. 2022. Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-Shot Cross-Dataset Transfer. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 3 (2022), 1623–1637. https://doi.org/10.1109/TPAMI.2020.3019967

Amir Semmo and Sebastian Pasewaldt. 2020. Graphite: Interactive Photo-to-Drawing Stylization on Mobile Devices. In *Proceedings SIGGRAPH Appy Hour* (Virtual Event, USA). ACM, New York. https://doi.org/10.1145/3388529.3407306

Jonathan Shade, Steven Gortler, Li-wei He, and Richard Szeliski. 1998. Layered Depth Images. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '98)*. Association for Computing Machinery, New York, NY, USA, 231–242. https://doi.org/10.1145/280814.280882

Meng-Li Shih, Shih-Yang Su, Johannes Kopf, and Jia-Bin Huang. 2020. 3D Photography Using Context-Aware Layered Depth Inpainting. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 8025–8035. https://doi.org/10.1109/CVPR42600.2020.00805