Deep Learning for Thin Structure Segmentation with an application to Image-Based Rendering

Masters 2 Internship (6 months)

George Drettakis, Inria Sophia Antipolis (France)
George.Drettakis@inria.fr

http://www-sop.inria.fr/members/George.Drettakis/

Context and goal

Image-Based Rendering (IBR) [Buehler 01, Chaurasia 13], allows free-viewpoint 3D navigation of a real scene, captured with a set of photographs ("multi-view dataset"). The pose of the cameras is estimated with Structure from Motion, and a coarse 3D model computed with Multi-view stereo (MVS); IBR reprojects the input photos into novel views, allowing highly realistic free-viewpoint navigation.

In recent work, we have developed novel algorithms that greatly improve the quality of IBR [Hedman 16, Hedman 18], including a solution [Thonat 18] that allows the treatment of the very hard case of thin structures, which typically cause MVS to fail (see Figure 1).

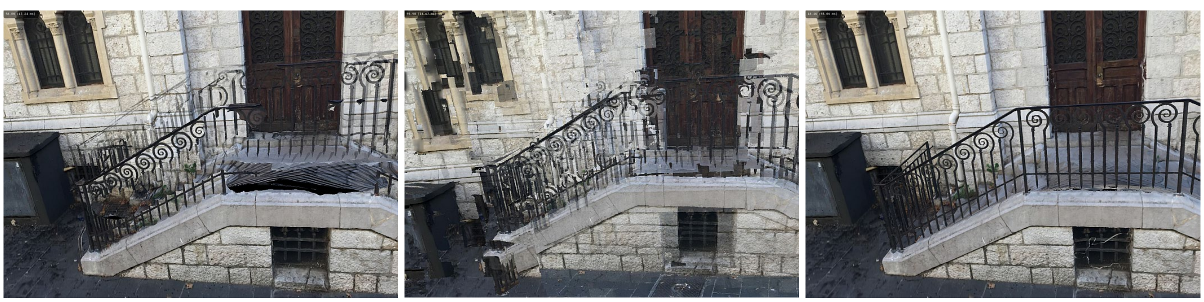

Figure 1: Thin structures are very hard to reconstruct causing traditional image-based rendering algorithms to fail. Left: [Buehler 01]; Middle [Ortiz-Cayon 15] Right: Our result [Thonat 18]

Our solution for thin structure segmentation and rendering [Thonat 18] assumes that the user manually defines the supporting geometry of the thin structure (currently planes and cylinders). Our goal in this internship is to find these structures automatically using learning-based solutions.

Approach

To solve this problem we will use deep learning with synthetic data for labelling. The first task in the internship will be the generation of the synthetic segmentation data, using a combination of purchased assets for thin structures and procedural generation. To automatically find the supporting geometry of thin structures, we will use two separate deep networks in an iterative configuration. The first network will identify thin structures in the images (a segmentation task, including multi-view constraints), while the second network will regress the parameters of the geometry supporting the thin structure (e.g., for planar structures, first estimating the normal, and then disambiguating depth by a sweep constrained by the surrounding MVS geometry).

We will first experiment with image-to-image translation CNNs (similar to [Isola 17]) and if these do not prove sufficient, we will investigate more sophisticated segmentation networks (e.g., [Chen 18]). The output of this segmentation will then be fed to a simple regression network, that will learn the parameters of the supporting geometry, again using synthetic ground truth data. The resulting supporting geometry and segmentation will first be used as initialization for [Thonat 18], but in a second step we will try and refine the learning approach to allow a complete end-to-end treatment of the problem.

The intern will develop this research in the existing C++/OpenGL software framework of the group that includes existing implementations of [Thonat 18], and a comprehensive framework for the generation of synthetic data.

Work environment and requirements

The internship will take place at Inria Sophia Antipolis, in the beautiful French Riviera. Inria will provide a monthly stipend between 450 and 1100€ depending on the situation of the candidate. The intern will work closely with the Ph.D. students in the group. Successful Masters internships may lead to a Ph.D. in the context of the ERC FUNGRAPH project (http://fungraph.inria.fr)

Candidates should have strong programming and mathematical skills as well as knowledge in computer graphics (a 4th year or higher graphics course is desirable), computer vision, geometry processing and machine learning.

References

[Buehler 01] C. BUEHLER, M. BOSSE, L. MCMILLAN, S. GORTLER, and M. COHEN. "Unstructured lumigraph rendering." In *Proceedings SIGGRAPH 2001*, pp. 425-432. ACM, 2001.
https://www.ics.uci.edu/~gopi/ICS280Win02/UnstructuredLumigraph.pdf

[Chaurasia 13] G. CHAURASIA, S. DUCHENE, O. SORKINE-HORNUNG, & G. DRETTAKIS. (2013). Depth synthesis and local warps for plausible image-based navigation. *ACM Transactions on Graphics (TOG)*, *32*(3), 30.
http://www-sop.inria.fr/reves/Basilic/2013/CDSD13/

[Hedman 16] P. HEDMAN, T. RITSCHEL, G. DRETTAKIS, G. BROSTOW, Scalable Inside-Out Image-Based Rendering, ACM Transactions on Graphics 35 (SIGGRAPH Asia), 6, December 2016
http://www-sop.inria.fr/reves/Basilic/2016/HRDB16/

[Hedman 18] P. HEDMAN, J. PHILIP, T. PRICE, J.-M. FRAHM, G. DRETTAKIS, G. BROSTOW, Deep Blending for Image-Based Rendering, ACM Transactions on Graphics 37 (SIGGRAPH Asia), 6, Dec. 2018
http://www-sop.inria.fr/reves/Basilic/2018/HPPFDB18/

[Isola 17] P. ISOLA, ZHU, J-Y, ZHOU, T., and EFROS, E. "Image-to-image translation with conditional adversarial networks." IEEE CVPR 2017.

http://openaccess.thecvf.com/content_cvpr_2017/papers/Isola_Image-To-Image_Translation_With_CVPR_2017_paper.pdf

[Chen 18] L. C., CHEN, G. PAPANDREOU, I. KOKKINOS, K. MURPHY, & A. L. YUILLE, (2018). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE transactions on pattern analysis and machine intelligence, 40(4), 834-848.

https://arxiv.org/pdf/1606.00915