High Quality Differentiable Rendering for Material Estimation in 3D Scenes

Masters 2 Internship (4-6 months)

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

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

Context and goal

Capturing 3D scenes — as opposed to isolated 3D objects — is an important challenge for Computer Graphics, avoiding the need to manually create scenes with geometry, materials and lights. An important aspect of capture is material acquisition; recent methods attempt to capture a material using a single photo, typically using a flash[Aittala 13, Li 17, Deschaintre 18 & 19]. The most recent such methods [Li 17, 18] including our own [Descaintre 18 & 19] use deep learning to infer the material properties, i.e., parameters of the Spatially Varying Bidirectional Reflectance Distribution Function (SVBRDF).

Most current methods are either limited to capturing a single patch [Deschaintre 18, 19; Lu 17, 18] or a single object [Nam 18, Li 18a]. Capturing more complex scene configurations poses numerous problems: dealing with occlusion boundaries, discontinuities in the surface parametrization and the placement and handling of more complex lights compared to the previous, simple patch/object and single flash configurations.

In parallel to the advances in material capture, recent work has seen the evolution of sophisticated differentiable renderers [Azinovic 19; Merlin 19], that can be directly used within deep learning frameworks such as pytorch.

Approach

In this internship, we will extend the approach of [Deschaintre 18,19] to the case of multi-view capture of a scene. We will start with entirely synthetic data, so the exact geometry is known, and we will use a UV-map or point-based representation of the 3D information of the scene. Instead of estimating simple maps that correspond to a single patch, we will use a similar CNN architecture to [Deschaintre 18,19] to estimate parts of the UV-map or point representation corresponding to each input image, allowing the estimation to happen in a single, shared space. We will use the differentiable renderer [Merlin 19] to estimate the rendered values in this shared space, as well as computing a fully rendered image for each input view given the current estimation of the maps during training.

We will start with simple scenes containing a single object, and compare the results to existing state of the art methods [Nam 18, Li18a]. We will then progressively increase the complexity of the scenes we treat, adding in occlusion boundaries and more complex lighting configurations. We expect that each level of complexity will require improvements to the network architecture, the loss and/or the training 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.

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. Successful Masters internships may lead to a Ph.D. in the context of the ERC FUNGRAPH project (http://fungraph.inria.fr)

References

[Azinovic 19] Azinović, Dejan, Tzu-Mao Li, Anton Kaplanyan, and Matthias Nießner. "Inverse Path Tracing for Joint Material and Lighting Estimation." *arXiv preprint arXiv:1903.07145* (2019).

[Aittala 13] Aittala, Miika, Tim Weyrich, and Jaakko Lehtinen. "Practical SVBRDF capture in the frequency domain." ACM Trans. Graph. 32.4 (2013): 110-1.

[Li 17] Li, Xiao, Yue Dong, Pieter Peers, and Xin Tong. "Modeling surface appearance from a single photograph using self-augmented convolutional neural networks." ACM Transactions on Graphics (TOG) 36, no. 4 (2017): 45.

[Li 18] Li, Z., Sunkavalli, K. and Chandraker, M., 2018. Materials for masses: SVBRDF acquisition with a single mobile phone image. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 72-87).

[Li 18a] Li, Zhengqin, Zexiang Xu, Ravi Ramamoorthi, Kalyan Sunkavalli, and Manmohan Chandraker. "Learning to reconstruct shape and spatially-varying reflectance from a single image." In *SIGGRAPH Asia 2018 Technical Papers*, p. 269. ACM, 2018.

[Deschaintre 18] Valentin Deschaintre, Miika Aittala, Frédo Durand, George Drettakis, Adrien Bousseau, Single-Image SVBRDF Capture with a Rendering-Aware Deep Network, ACM Transactions on Graphics (SIGGRAPH Conference Proceedings), Volume 37, Number 128, pages *15*, Aug 2018 http://www-sop.inria.fr/reves/Basilic/2018/DADDB18/

[Deschaintre 19] Valentin Deschaintre, Miika Aittala, Frédo Durand, George Drettakis, Adrien Bousseau, Flexible SVBRDF Capture with a Multi-Image Deep Network, Computer Graphics Forum (Proceedings of the Eurographics Symposium on Rendering), Volume 38, Number 4, July 2019 http://www-sop.inria.fr/reves/Basilic/2019/DADDB19/

[Nam 18] Nam, G., Lee, J.H., Gutierrez, D. and Kim, M.H., 2018, December. Practical SVBRDF acquisition of 3D objects with unstructured flash photography. In *SIGGRAPH Asia 2018* Technical *Papers* (p. 267). ACM.

[Merlin 19] Merlin Nimier-David, Delio Vicini, Tizian Zeltner, and Wenzel Jakob. 2019. Mitsuba 2: A Retargetable Forward and Inverse Renderer. In *Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 38(6). http://rgl.epfl.ch/publications/NimierDavidVicini2019Mitsuba2