PhD THESIS

LIGHTING ESTIMATION FROM IMAGES FOR SEAMLESS INTEGRATION OF VIRTUAL OBJECTS INTO REAL SCENES



Figure 1: Two qualitative examples showcasing DepthLight's capability to automatically insert virtual bunnies taking into account spatial-aware lighting effects from a single low dynamic range and limited field-of-view image.

TEAM

VIRTUS, INRIA Center at Rennes University

KEYWORDS

Virtual Production, Image-Based Rendering, Computer Graphics, Real-Time Visual Effects

CONTEXT

All visual effects used on feature films have a process where virtual objects, whether creatures or set extensions, require to have a perfect synthetic lighting matching the real on-set lighting. If they fail to capture or reproduce the real lighting, the human eye is trained enough to instantly detect issues and classify the result as being not realistic. Even today on most recent super-productions, that lighting seems wrong for some situations. This is mostly due the inherent technique used -the HDRI chrome and diffuse balls- that fails to capture the local lighting changes. These are just a coarse approximation of the real lighting and they do not take into account occlusions and interreflections. To correct this, some VFX companies will have artists to manually correct the lighting which may induce additional flaws in the visual integration of virtual objects.

Past year, a preliminary work has been achieved using single images as input. It allowed us to estimate the global illumination of a real scene. The first step computed an HDR panorama from a single image [2], whilst an approximation of the 3D model of the scene was obtained using AnyDepth [3]. By reprojecting the estimated HDR panorama onto the 3D model, we generate a first approximation of the global illumination and thus insert virtual objects into real scenes in a photorealistic and visually plausible way, using Unreal Engine [1]. An illustration of that technique is shown in Figure 1.

OBJECTIVES

The objective of this thesis is to work on new techniques for estimating the position, orientation and intensity of light sources from videos, to seamlessly integrate virtual objects into real scenes.

We propose here to build on this work by first improving the current limitations and then generalizing the approach. The first step is to be able to work from videos where the lighting is dynamic. Indeed,

the current approach only allows static lighting as we use a single image to compute the lighting. The key challenge lays in the ability to generate temporally stable light estimations over the sequence. A second step is to design a validation framework to assess the preciseness of the light reconstruction and the quality of the generated results, typically by using the ground truth generated by a game engine such as Unreal Engine (UE). More precisely, we would use a photo realistic scene and use not only the real video created by UE but also the 3D model and the HDR representation of the scene to compare our results to the same simulation in UE. Additional extensions would include the possibility to have a multi-pass rendering approach where the synthetic objects illumination also impacts the real scenes (consider for example the color bleeding of a virtual red sphere over a real white wall).

In addition to the previous research directions, to improve the generalization capacity the thesis will also focus on exploring techniques using radiance scene representations such as 3D Gaussian Splatting to model more precise light behaviors than emissive meshes [5] and also explore how diffusion-based techniques could be adapted to handle light sources estimation, in the way ray diffusion is used to perform camera pose estimation [6].

The application of the results to real scenes will be then investigated, either from single images or 360 videos provided by VFX companies and movie productions.

The intern will have the opportunity to work with Sam Boivin senior INRIA researcher and former 8-years CTO of the world-leading real-time AR company for VFX ncam-technologies [4], and Marc Christie, associate professor at University of Rennes.

REFERENCES

[1] https://www.unrealengine.com/
[2] Wang, Jionghao & Chen, Ziyu & Ling, Jun & Xie, Rong & Song, Li. (2023). 360-Degree Panorama Generation from Few Unregistered NFoV Images. 10.48550/arXiv.2308.14686.
[3] Yang, Lihe & Kang, Bingyi & Huang, Zilong & Zhao, Zhen & Xu, Xiaogang & Feng, Jiashi & Zhao, Hengshuang. (2024). Depth Anything V2. 10.48550/arXiv.2406.09414.
[4] https://www.ncam-tech.com/
[5] Bernhard Kerbl and Georgios Kopanas and Thomas Leimkühler and George Drettakis. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. ACM Transactions on Graphics, volume 42(4), July 2023
[6] Zhang, Jason Y and Lin, Amy and Kumar, Moneish and Yang, Tzu-Hsuan and Ramanan, Deva and Tulsiani, Shubham . Cameras as Rays: Pose Estimation via Ray Diffusion. International Conference on Learning Representations (ICLR), 2024.

SUPERVISION

Sam Boivin, INRIA Researcher

Marc Christie, Associate Professor, University of Rennes 1

CONTACT

samuel.boivin@inria.fr

marc.christie@irisa.fr