HAL
open science

Demonstrating COLIBRI VR, an Open-Source Toolkit to Render Real-World Scenes in Virtual Reality

Grégoire Dupont de Dinechin, Alexis Paljic

To cite this version:

Grégoire Dupont de Dinechin, Alexis Paljic. Demonstrating COLIBRI VR, an Open-Source Toolkit to Render Real-World Scenes in Virtual Reality. 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Mar 2020, Atlanta, Georgia, United States. hal-02492733

HAL Id: hal-02492733

https://hal.science/hal-02492733v1

Submitted on 27 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Demonstrating COLIBRI VR, an Open-Source Toolkit to Render Real-World Scenes in Virtual Reality

Grégoire Dupont de Dinechin* Alexis Paljic†

Centre for Robotics, MINES ParisTech, PSL University - Paris, France

ABSTRACT

This demonstration showcases an open-source toolkit we developed in the Unity game engine to enable authors to render real-world photographs in virtual reality (VR) with motion parallax and view-dependent highlights. First, we illustrate the toolset's capabilities by using it to display interactive, photorealistic renderings of a museum's mineral collection. Then, we invite audience members to be rendered in VR using our toolkit, thus providing a live, behind-the-scenes look at the process.

Index Terms: Computing Methodologies—Computer Graphics—Graphics Systems and Interfaces—Virtual Reality; Computing Methodologies—Computer Graphics—Image Manipulation—Image-Based Rendering;

1 INTRODUCTION

1.1 Motivations

Many immersive experiences rely on virtual environments and assets created from sets of real-world photographs. Indeed, images and videos are a convenient way of capturing the real world's scenery for VR, e.g. to create immersive virtual visits, high-fidelity training environments, and vivid entertainment experiences. These *image-based VR* experiences are made even more compelling when the photography-based environments and assets provide *motion parallax* - i.e. the sensation that, when moving one's head, closer objects move faster than objects further away - and render *specular highlights* - i.e. convincingly expose bright reflections on specular surfaces by adapting objects' colors based on the viewpoint from which they are seen (see Fig. 1). Multiple recent works have thus demonstrated how to achieve motion parallax and view-dependent effects for image-based VR scenes rendered from both 360° [3,6] and perspective [2,4] image datasets, by estimating underlying geometry and applying different image-based rendering techniques.

However, despite this interest for creating virtual reality content from photography-based assets and environments, there seem to be very few integrated, open-source interfaces dedicated to providing easy access to image-based modeling and rendering tools. Indeed, although 3D reconstruction is now commonly used by VR content creators, e.g. using consumer photogrammetry toolkits, the reconstruction workflow is generally unadapted to rapid prototyping by small research teams, as it is often costly and complex. As for view-dependent rendering, its study and use seems to be mostly confined to an expert segment of the graphics community, and it is only rarely examined in the context of VR user experience research: instead of dynamically updating the reconstructed scene's color from the closest images to render specular highlights, color data is commonly averaged into static texture maps during reconstruction, which is likely to result in lower-fidelity output views at render time.

*e-mail: gregoire.dupont_de_dinechin@mines-paristech.fr
†e-mail: alexis.paljic@mines-paristech.fr

©2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

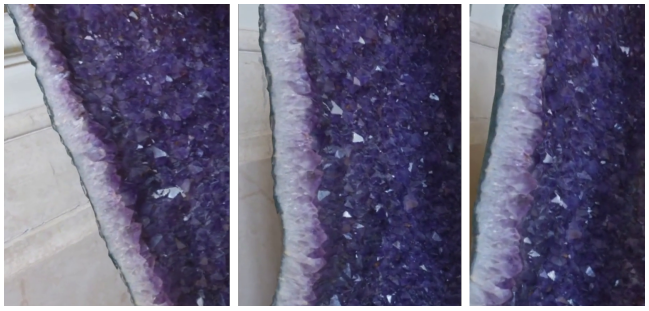

Figure 1: Close-ups of an amethyst rendered using COLIBRI VR. Because we use a view-dependent rendering method, specular highlights appear on the mineral's facets as the viewpoint changes.

1.2 Our approach

To help tackle this issue, we developed the Core Open Lab on Image-Based Rendering Innovation for Virtual Reality (COLIBRI VR), an open-source, integrated resource that provides image-based rendering methods adapted for VR, based on which to discuss, prototype, and lead research related to user experience in photography-based virtual environments. Because the toolset and codebase are openly-available, authors wanting to learn more about how to implement image-based modeling and rendering algorithms now have concrete examples to work with, that they can try out on their data and eventually improve. This also means that results obtained using the toolkit can more easily be compared with and replicated by other research teams. Furthermore, because COLIBRI VR is integrated as an interface within the Unity game engine, it can be used as a natural addition to VR content creators' usual development pipeline, to develop interactive experiences that combine synthetic virtual objects and photography-based assets. By linking the toolset to open-source software for 3D reconstruction and mesh simplification such as COLMAP [5] and Blender, we also provide an easy way for authors to estimate geometry from sets of images from within Unity.

In terms of architecture, COLIBRI VR essentially consists of C# scripts and a series of ShaderLab (Cg/HLSL) vertex/fragment and compute shaders. The first define an overarching framework for the different processing methods and provide the toolkit with a graphical user interface. With the second, we implement modeling and rendering operations, such as mesh generation from depth maps and view-dependent blending: we thus provide implementations of several core image-based rendering algorithms, such as Unstructured Lumigraph Rendering [1]. Because we target interactive virtual reality, one of the challenges during development was to adapt and optimize these algorithms so as to maintain high framerates, in order not to cause discomfort. For the same reason, we had to regularly check that the implemented shaders and scripts effectively enable immersed users to act on the image-based assets in a variety of ways, e.g. grab them, change them to be larger or smaller, or make them interact with synthetic 3D objects, as we expect such interactions to be useful in many typical image-based VR use cases.

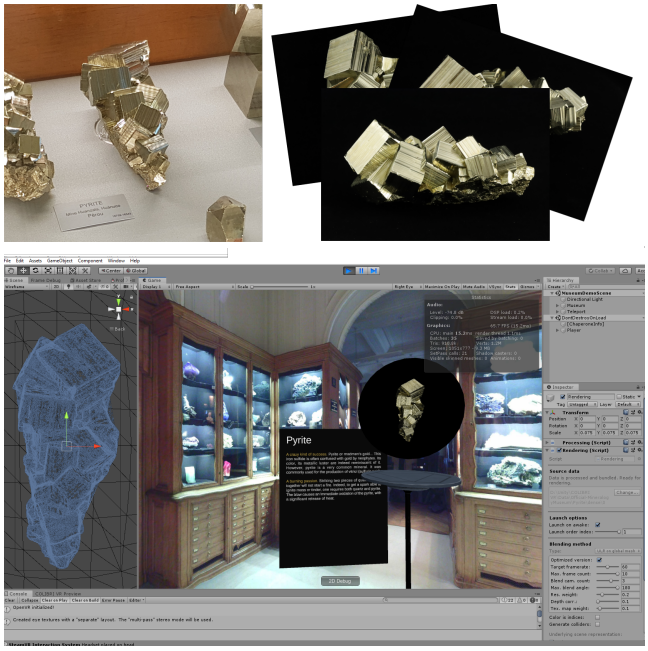

Figure 2: Using COLIBRI VR to render minerals as interactive assets in a VR museum. Top: a pyrite in its museum stand, and photographs of the pyrite used as input for the process. Bottom: the rendered 3D mineral and the toolkit's interface in the Unity game engine.

2 Research Demonstration

2.1 An interactive VR mineralogy museum

The first segment of our demonstration consists in inviting audience members to try out an interactive VR visualization of the MINES ParisTech Mineralogy Museum's collection of minerals (see Fig. 2). The content for this segment is pre-processed before the demonstration: we designed a virtual environment in which to display the virtualized minerals, and used COLIBRI VR's interface to transform the input photographs of these minerals into scene representations adapted for real-time view-dependent rendering in VR. When trying out the experience, audience members are thus able to examine the minerals in 3D space by walking around them and picking them up, and can observe the specular reflections captured as light bounced off the minerals' facets during acquisition.

The objective, for the museum's curators, is to display the collection interactively and with high fidelity in an online showcase, so as to give far-away viewers a teaser of the physical visit and encourage them to come in person. The experience may also enable users to more closely examine the displayed pieces: for instance, many of the real-world minerals cannot be touched due to being fragile and costly, and may occasionally be unavailable for viewing; the virtualized minerals therefore provide a complementary learning experience, as they can be taken and observed from up close, no matter the current state or location of their real-world counterpart. This segment is also an opportunity to showcase COLIBRI VR's potential in the context of a specific use case. Indeed, the specular highlights appearing on a mineral's facets are an important visual cue for understanding its visual appearance and material properties: without the ability to render these highlights, many minerals thus could not effectively be rendered as 3D assets in a virtual museum.

2.2 Live reconstruction of audience members with view-dependent gaze following

The second segment of our demonstration aims to illustrate a custom view-dependent rendering effect enabled by the toolkit, by which

Figure 3: Using COLIBRI VR to render custom view-dependent effects, such as having a virtualized character's gaze follow the user.

a virtualized model of a person will seem to always be looking towards the user in VR (see Fig. 3). This effect is inspired by the *Welcome to Light Fields* [4] demonstration, and is obtained by taking multiple pictures of a participant from different angles, in which the participant stays still but keeps looking towards the camera as it moves to different viewpoints. Using the graphical user interface, we call COLMAP to reconstruct a 3D model from these photographs, and use COLIBRI VR's implementation of view-dependent rendering methods to render the model. By dynamically selecting the images most relevant for the current viewpoint before blending them to create the output view, we thus select images taken from a viewpoint close to the user's current one, which are therefore images in which the person seems to be looking towards the user. In this way, the rendered person's gaze will therefore seem to follow the immersed users as they move around the virtual space.

This segment is an opportunity to answer the audience's questions and present COLIBRI VR's processing and rendering interface in more detail. Because the reconstruction operations are not real-time, we use one computer to render the museum demonstration for the audience, and another to process the photographs and prepare them for rendering. Once the data is ready, we then show the rendered participant in VR inside the head-mounted display, before starting the demonstration anew with a different group of audience members.

Acknowledgments

The authors wish to thank Eloïse Gaillou and Didier Nectoux from the MINES ParisTech Mineralogy Museum for their support on the virtual reality mineralogy museum project.

References

[1] C. Buehler, M. Bosse, L. McMillan, S. Gortler, and M. Cohen. Unstructured lumigraph rendering. In *28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, p. 425–432. ACM Press, 2001. doi: 10.1145/383259.383309

[2] C.-F. Chen and E. S. Rosenberg. Automatic generation of dynamically relightable virtual objects with consumer-grade depth cameras. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1295–1296. IEEE, Mar. 2019. doi: 10.1109/VR.2019.8798037

[3] J. Huang, Z. Chen, D. Ceylan, and H. Jin. 6-DOF VR videos with a single 360-camera. In *2017 IEEE Virtual Reality (VR)*, pp. 37–44. IEEE, Mar. 2017. doi: 10.1109/VR.2017.7892229

[4] R. S. Overbeck, D. Erickson, D. Evangelakos, M. Pharr, and P. Debevec. A system for acquiring, processing, and rendering panoramic light field stills for virtual reality. *ACM Transactions on Graphics (TOG)*, 37(6), Dec. 2018. doi: 10.1145/3272127.3275031

[5] J. L. Schönberger and J.-M. Frahm. Structure-from-motion revisited. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4104–4113. IEEE, June 2016. doi: 10.1109/CVPR.2016.445

[6] A. Serrano, I. Kim, Z. Chen, S. DiVerdi, D. Gutierrez, A. Hertzmann, and B. Masia. Motion parallax for 360° RGBD video. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):1817–1827, May 2019. doi: 10.1109/TVCG.2019.2898757